INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International
300 N. ZEEB RD., ANN ARBOR, MI 48106

Dush, David Michael

A META-ANALYSIS OF COGNITIVE RESTRUCTURING THERAPY

Kent State University PH.D. 1981

A Meta-Analysis of Cognitive Restructuring Therapy

A dissertation submitted to the
Kent State University Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

by

David Michael Dush

December, 1981

Dissertation written by

David Michael Dush

B.Sc., Michigan State University, 1976

M.A., University of Manitoba, 1978

Ph.D., Kent State University, 1981

Approved by

Michael Hunt, Chairman, Doctoral Dissertation Committee

Roy L. Lilly, Members, Doctoral Dissertation Committee

Harold G. Schutte,

John Ahammer,

[illegible],

Accepted by

Roy L. Lilly, Chairman, Department of Psychology

Robert Rasell, Dean, Graduate College

Acknowledgments

I am deeply grateful to all of those who helped in this endeavor with their time, effort, and encouragement. Special thanks go to Dr. Michael Hirt. As my advisor and chairman of the dissertation committee, his consultation, high standards, and patient but persistent encouragement surely account in large part for strengths the manuscript may have.

I am also indebted to the members of my dissertation committee. Drs. Roy Lily, Harold Shroeder, and Sonia Blixt were more than generous with their time, and with their critieal contributions were very helpful in solidifying the aims of the study. Thanks also go to Debby Rich, a very competent and enthusiastic assistant who made my tasks much easier. Kathy Miller was invaluable in her assistance with the manuscript. The critical readings and encouragement of Billie Hothem were especially appreciated. This work was partially supported by a grant from the Graduate College, which helped significantly in attaining the present scope. The interlibrary loan personnel also graciously spent many hours on my behalf.

Finally, I owe the deepest acknowledgment to my parents, George and Hellen Dush, for their many years of uninterrupted support, understanding, and encouragement. It is to them that this endeavor is dedicated.

Table of Contents

List of Figures

Figure Page

List of Tables

A Meta-Analysis of Cognitive Restructuring Therapy

With the increasing rate at which scientific research produces new data and developments, researchers face the nearly impossible task of keeping abreast of the literature of their specialty. Kulik, Kulik, and Cohen (1979) suggest that, in most specialties, a researcher devoting half of each day to reading will at best cover only about 10% of the new work in the area. By the year 2000 the number of articles of relevance is expected to increase fourfold.

Clearly there is a substantial and growing need for reviewing, summarizing, and integrating scientific findings. In a discussion of the techniques and issues involved in integrating research, Feldman (1971) elaborated on several areas in which integration can be of value: (a) As a "compendium of research," integration gathers, condenses, and abstracts previously dispersed information; (b) Bibliographic listings of research and theory are provided which highlight the amount of work which has or hasn't been done in various areas, and where the deficits or surfeits lie; (c) Large, broad based samples of investigations partially simulate a "national sample," contributing to generalizing beyond local or situational variables; (d) Comparisons of studies which reasonably closely

duplicate one another in important aspects can simulate replication; (e) "Secondary analysis" can address new or corroborative hypotheses by performing different statistical operations on the data or parameters reported in the primary investigations; (f) Old or newly generated hypotheses can be evaluated, and theoretical issues can be generated for further research.

Kulik et al. (1979) suggest that the difficulties of summarizing research are especially troublesome in applied social science:

> Each applied study has its own wrinkles, and the results of applied studies often vary in confusing ways from setting to setting. Although a few replications can often establish an experimental phenomenon in the exact sciences, applied investigators often produce dozens of studies without establishing a clear trend. (p. 307)

With reference tc psychotherapy research in particular, Smith and Glass (1977) have reached the conclusion that:

> Scholars and clinicians are in the rather embarrassing position of knowing less than has been proven, because knowledge, atomized and sprayed across a vast landscape of journals, books, and reports, has not been accessible. (p. 760)

The traditional medium for integrating research--the "literary" or "narrative" review--has been criticized on several counts. First, out of practicality, the literary review typi-

cally concentrates on much less than the total population of reports in a given area, leaving the representativeness of the sample uncertain. Second, Smith and Glass have pointed out that the "voting method" (counting the number of studies with outcomes favoring one hypothesis over another)"is too weak to answer many important questions and is biased in favor of large-sample studies" (1977, p. 752). Third, as progressively more complexity is dealt with in the literary review, the conclusions become less straightforward and are more susceptable to biases or idiosyncrasies of the reviewer. As Glass more forcefully suggests:

> A common method for integrating several studies with inconsistent findings is to carp on the design or analysis deficiences of all but a few studies--those remaining frequently being one's own work or that of one's students or friends--and then advance the one or two 'acceptable' studies as the truth of the matter. (1976, p. 4)

Lastly, the size of the effects of relationships are rarely considered, and the conclusions may be inappropriately weighted with regard to sample size or other parameters of the body of research (Cooper, 1979).

Meta-Analysis

Glass (1976) has proposed a direct extension of the principles of statistical analysis to the problems of integra-

ting or aggregating data from several sources. He distinguishes three levels of statistical analysis: (a) primary analysis, the first analysis of the data in an investigation as reported; (b) secondary analysis, which refers to the reanalysis of the data from a given study, ordinarily as a check or to explore variations in the analytic approach; and (c) meta-analysis, which refers to "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings" (Glass, 1976, p. 3). Glass cogently argues that the need for more thorough integration of knowledge already accumulated is an urgent one, and that meta-analysis:

> Connotes a rigorous alternative to the casual narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding research literature. (1976, p. 3)

Meta-analysis is clearly distinct from the literary review in its approach to integration of findings. Once certain assumptions and decisions about organizing and scaling the data are made, the meta-analysis of the data represented in the body of studies under scrutiny proceeds in essentially the typical fashion of statistical analysis. It is broadly applicable to any sufficiently large and representative group of studies which are related in that "(a) they share a common conceptual hypothesis or (b) they share operations for the realization of the independent or dependent variables,

regardless of the conceptual focus" (Cooper, 1979, p. 133).

The starting point in meta-analytics is the expression of the data of interest in each study in a metric common across studies. Once the focal data in each study are expressed in a format and scale common across all studies, statistical (meta-) analyses can proceed which (a) treat the results of individual studies as the units of analysis, and (b) encompass the findings of all studies (or meaningful subsets) within the scope of the aggregative analysis.

For example, in outcome studies, where one or more "treated" groups are contrasted with one or more groups receiving varied or no treatment on some dependent variable(s) of interest, each outcome can be expressed as an effect size (Cohen, 1977). That is, the influence of any or all treatments in a given study on an outcome variable measured in the study can be meaningfully represented as the mean difference between the treated group(s) and a control group on the outcome variable--a representation format Cohen has labelled effect size. The effect size can then be converted to a standarized scale, by dividing the mean difference between groups by an estimate of the variability of untreated population (Glass, 1978). Thus, each treatment effect in the studies under review can be represented in a comparable format and expressed in a common, standarized scale, accounting for variability observed in the absence of treatment. Expressed as a formula:

$$ES_{t-c} = \frac{\bar{X}_t - \bar{X}_c}{SD_c}$$

where $\bar{X}_t$ and $\bar{X}_c$ refer to the means of the treatment group (t) and control group (c) on the dependent variable in question, SD_c is the standard deviation of the control group on the dependent variable, and ES_{t-c} refers to the standardized effect size of treatment t.

The key conversion of results across studies thus involves expression of each treatment effect as an effect size. There are no inherent constraints on the variability in number or nature of groups or outcome variables. Once comparably expressed the data lend themselves to statistical analyses ranging in focus from basic aggregative summary and description to complex correlational analysis and inference. Flexibility is also realized in the level of aggregation, as statements can be made with reference to either (a) all effects on all outcome measures across all studies under analysis, simultaneously, or (b) any meaningful subdivision of ES's by treatments (e.g., therapy performed by the most experienced therapists), outcome measures (e.g., only effects of treatment on self-concept), or by studies (e.g., only studies which used matched comparison groups).

Illustrative points can be drawn from the meta-analysis of 375 studies of the outcome of psychotherapy (Smith & Glass, 1977). Whereas a long history of literary reviews drawing

on different samples of outcome studies and different interpretations of the same studies had failed to approach consensus regarding the effectiveness of psychotherapy, Smith and Glass conducted a more objective meta-analysis which aggregated the findings from all studies located which compared at least one therapy treatment group to an untreated or different therapy group. At the highest level of aggregation it was demonstrated that, across all types of therapy and outcomes across all 375 studies (reflecting 833 ES's) the average treated client was better off than 75% of the untreated controls. Various partitionings of the data provided for demonstrations of the moderating effects of other characteristics of the data, such as type or class of therapy, quality of research design, adequacy of measurement, characteristics of subjects and therapists, etc., on ES. Clearly, a wide range of debated issues in the controversy over psychotherapy effectiveness were addressed from a statistical perspective.

Other recent investigations have similarly applied meta-analysis to a wide range of data and a variety of research questions. White (1979) found only about 5% of the variability in academic achievement predictable from socioeconomic status, while literature cited suggested that a much higher degree of relationship had long been thought to have been established. His findings were based on a meta-analysis of 100 studies. In another meta-analysis of 75 studies comparing personalized systems of instruction (PSI) to conventional scholastic

instructional methods, Kulik et al. (1979) found that PSI raised the final exam score of the average student from the 50th to the 70th percentile. Similarly, additional recent meta -analyses have addressed such varied areas as experimenter effects (Rosenthal, 1976), sex differences in conformity research (Cooper, 1979), the effect of class size on achievement (Glass & Smith, 1979), the effects of television on social behavior (Hearold, 1979), the comparative effects of drug therapy and psychotherapy (Miller, 1979), and sex bias in psychotherapy (Smith, 1980).

The notions inherent in meta-analysis are not new, having their roots in work on obtaining overall probability levels across studies (Fisher, 1938; Pearson, 1938). Meta-analysis in the present sense has only recently been advanced, however. Light and Smith (1971) addressed the technical issues of aggregating research findings, but the criteria for studies to be suitable for their proposed procedures--including direct access to the original raw data--left the applicability too severely restricted for most purposes in applied research. The most frequently applied procedures for meta-analysis in the present sense originated with Glass (1976, 1978) and his colleagues (Glass & Smith, 1979; Smith & Glass, 1977; Smith, Glass, & Miller, 1980), although there are current additional, varied meta-analytic procedures that differ procedurally and according to the nature of the data and research questions to be addressed (Barton & Glass, 1979; Glass, 1978; Hunter,

1979; Rosenthal, 1976, 1978).

The advantages and distinctions of meta-analysis in relation to conventional methods of reviewing and integrating are several. First, as a statistical procedure, meta-analysis offers greater rigor and objectivity than the traditional literary review. Second is the potential for greater clarity and parsimony in providing a summary of the mutiplicity of characteristics and qualifications in the typical large body of research. As Glass notes, the results of hundreds of studies "can no more be grasped in our traditional narrative than one can grasp the sense of 500 test scores without the aid of techniques for organizing, depicting, and interpreting data" (1976, p. 4). Third, under certain circumstances, statistical inferences can be made about interpretations drawn from the relationships addressed in the meta-analysis. Often, however, conventional statistical analyses are inappropriate or meaningless, especially with very large groups of studies, for reasons which will be discussed later. Fourth, basing conclusions on data aggregated across a large number of studies avoids many of the pitfalls of aggrandizing the statistical significance levels of individual studies or being misled by the sampling error reflected in individual studies (Hunter, 1979). Fifth, the very methodological discrepancies or weaknesses which typically obscure or weaken the conclusions from literary reviews can be quantitatively dealt with in meta-analysis. Design characteristics (e.g., random assignment

vs. matching), appropriateness of statistical analyses, and so forth, can be coded and treated as meta-analysis covariates, to evaluate or to "partial out" their impact on ES, the dependent variable. Sixth, other variables or combinations of variables of substantive interest can be similarly evaluated objectively as covariates, typically by statistically regressing them on the effect sizes. Hence, Smith and Glass (1977) found significant relationships between psychotherapy outcome and variables such as client I.Q. and similarity of clients and therapists. Seventh, meta-analysis offers greater precision in that the calculation of ES's provides an exact index of the magnitude and direction of treatment effects, and consequently is a more powerful technique for accurately detecting true relationships than the cruder alternatives available to the narrative method, such as the voting method. Finally, studies can easily be weighted to statistically account for differences in sample size, number of outcome measures, etc.

In summary, it has been suggested that in many respects, the accumulation of research in the social sciences has overshadowed its assimilation. The traditional literary review, by nature of its inherent limitations, has not provided a medium for adequate integration and communication of research when a large number of studies and inconsistencies are involved. There is an increased liklihood of repeating old mistakes, rediscovering previously demonstrated phenomena,

and attenuating new advancements by building on less than our established foundation of scientific knowledge. Meta-analysis has been presented as a distinct methodology which appears to overcome many of the limitations of the literary review and is particularly suited to dealing, in a systematic fashion, with the large numbers of varied studies typically found in the social sciences.

Cognitive Behavior Therapy

As a prospect for application of meta-analysis, the literature dealing with cognitive behavior therapy seems appropriate on several counts. On the one hand, cognitive behavior therapy--a term presently used to encompass a wide range of techniques variously referred to as cognitive therapy, cognitive behavior modification, cognitive learning therapy, etc.--has achieved a recent popularity that fosters a high and increasing rate of related theoretical and empirical research (Ledwidge, 1978). Perhaps in part a function of recency and proliferation, however, no integrative review has appeared that attempts to summarize or analyze the relevant research to an exhaustive degree. Integration is found to be more selective, in theory-oriented work (e.g., Beck, 1976; Meichenbaum, 1977), as a selective part of other review (such as the inclusion of Rational-Emotive Therapy (RET) research in the meta-analysis by Glass and Smith (1977)), or in work addressing a specific aspect, such as the comparability of

cognitive and behavioral techniques (Ledwidge, 1978). To date there appears to be no comprehensive review, literary or meta-analytic, which represents an exhaustive integration of the findings of controlled empirical research in the area. In light of multiplicity of techniques being employed (Mahoney & Arnkoff, 1978), the proliferation of new studies, and concerns voiced by some (e.g. Ledwidge, 1978) that the movement may be a "step in the wrong direction," the lack of a thorough integration of the empirical evidence already accumulated becomes a critical concern. It is to this void that the present investigation is addressed.

An immediate and key concern in addressing the literature of cognitive behavior therapy lies in attempting to more clearly define the domain implied. It is difficult and somewhat arbitrary to delimit boundaries separating "cognitive" from "behavioral" therapies, or even "more cognitive" from "less cognitive" techniques. This confronts inherent issues of the existence and definition of "cognitions," and the importance of cognitions in behavior therapy, other therapies, and behavior change in general, which have long been debated and certainly will not be resolved in the present analysis. Yet, some form of topography of operations is helpful in delimiting the focus of a meta-analysis. Mahoney and Arnkoff (1978) have recently summarized the various techniques, procedures, and frameworks that are generally considered to fall under the broad categories of "cognitive" and "self-control"

therapies. Their analysis suggested that the various approaches could be more usefully categorized in a conceptually meaningful way. Several techniques which appear to share certain fundamental assumptions were referred to as "Cognitive Learning Therapies," which, in turn, were further divided into three subordinate classes: (a) Cognitive Restructuring, (b) Coping Skills Therapy, and (c) Problem Solving Therapy. The resultant classification scheme, and the contemporary procedures included, are represented in Table 1.

While the variations of Cognitive Learning Therapy still differ in historical roots, operations, and other important ways, Mahoney and Arnkoff argue that there is some communality of basic assumptions, including:

> 1. Humans develop adpative and maladaptive behavior and affective patterns through cognitive processes (selective attention, symbolic coding, etc.).
> 2. These cognitive processes can be funtionally activated by procedures that are generally isomorphic with those of the human learning laboratory (although there may be other procedures which activate the cognitive processes as well).
> 3. The resultant task of the therapist is that of a diagnostician-educator who assesses maladaptive cognitve processes and subsequently arranges learning experiences that will alter cognitions and the behavior and affective patterns with which they correlate. (1978, p. 692)

Table 1

Contemporary Cognitive Learning Therapies[a]

Cognitive Restructuring	Coping Skills Therapies	Problem Solving Therapies
Rational-Emotive Therapy (Ellis, 1962)	Covert Modeling (Cautela, 1971)	Behavioral Problem Solving (D'Zurilla & Goldfried, 1971)
Self-Instruction (Meichenbaum, 1977)	Coping Skills Training (Goldfried, 1971)	Problem-Solving Therapy (Spivack & Shure, 1974)
Cognitive Therapy (Beck, 1976)	Anxiety Management (Suinn & Richardson, 1971)	Personal Science (Mahoney, 1974)
	Stress Inoculation (Meichenbaum, 1977)	

[a]Based on a table presented by Mahoney and Arnkoff (1978, p. 703).

Certain therapies and procedures which might easily be considered "cognitive" to some degree are not considered cognitive learning therapies within this scheme. Two such groups are: (a) behavioral self-control procedures (e.g. stimulus control, self-monitoring, self-reward, self-punishment); and (b) covert conditioning procedures (e.g., covert counterconditioning, thought stopping, coverant control, covert sensitization, covert reinforcement, covert extinction, covert response cost).

The exclusion of the self-control literature is largely due to the primary or exclusive focus of the techniques on behavioral strategies rather than cognitive processes, which seems to set them apart as a group. The covert conditioning literature is distinguished partly because, while the various procedures within that group are quite similar in orientation, they vary considerably from most techniques in the cognitive learning therapy group. Additionally, the current status of research from the covert condition perspective is not clearly supportive. Mahoney and Arnkoff conclude that:

> 1. There is relatively little research on the processes and effects of covert conditioning therapies.
> 2. What research does exist is generally poor in either internal or external validity.
> 3. Those few studies with more adequate experimental methodologies have generally lent little support to the clinical promise of covert conditioning therapies. (1978, p 710)

Concerns in delimiting a focus for the present meta-analysis were governed by the reconciliation of two objectives: (a) maintenance of sufficient breadth to allow for an integration which would address the major issues across the cognitive behavioral therapies, and (b) identification of a body of research which was meaningfully cohesive in basic conceptual orientations or operationalizations. Although meta-analysis can, as noted, address high levels of aggregation across many sources of variation, a tradeoff is implied. Given limited resources and data, there are advantages in beginning with a somewhat more homogeneous, yet substantial body of research, and attending to concerns at finer levels of distinction which apply to that circumscribed domain.

In a preliminary search of the cognitive behavior therapy literature it became apparent that the majority of controlled outcome studies include techniques within Mahoney and Arnkoff's category of cognitive restructuring. As the category implies, all three orientations share similar procedures and a singular goal, the modification of maladaptive thought patterns.

In applications of cognitive restructuring procedures, various supplemental techniques are often incorporated, such as self-monitoring, homework assignments, relaxation, etc. Across most applications, however, cognitive restructuring interventions are defined by two dimensions: (a) rational restructuring and (b) self-instructional training. Rational

restructuring refers to the process of recognizing maladaptive cognitions and, via some dialog with the therapist, endeavoring to replace them with more adaptive cognitions. Thus, with RET (Ellis, 1962), therapists would refute irrational beliefs and foster rational ones; with Cognitive Therapy (Beck, 1976), therapists would illucidate and correct inaccuracies and distortions of thought patterns; and with Self-instructional Therapy (SIT) (Meichenbaum, 1977), therapists would identify maladaptive self-statements and encourage their replacement with adaptive ones.

A second procedural dimension is the direct modification of self-statements. At the extremes, some applications of RET make no explicit use of this technique and some applications of SIT use it to the exclusion of explicit rational restructuring. In most applications both are present; Still, on the average, RET makes much greater use of rational restructuring, SIT makes much greater use of self-statement modification, and Cognitive Therapy lies somewhere between (see Figure 1).

Within the range of therapies most typically considered cognitive, it appears that a focus on cognitive restructuring therapies addresses a relatively homogeneous group of procedures which represent a reasonable compromise of breadth and precision for the present meta-analysis. A large proportion of the relevant empirical research is encompassed by this focus. RET studies were not presently included, since they have been subject to meta-analysis elsewhere (Smith et al., 1980). To provide selection criteria further assuring homegeniety, all other controlled, clinical outcome studies of cognitive

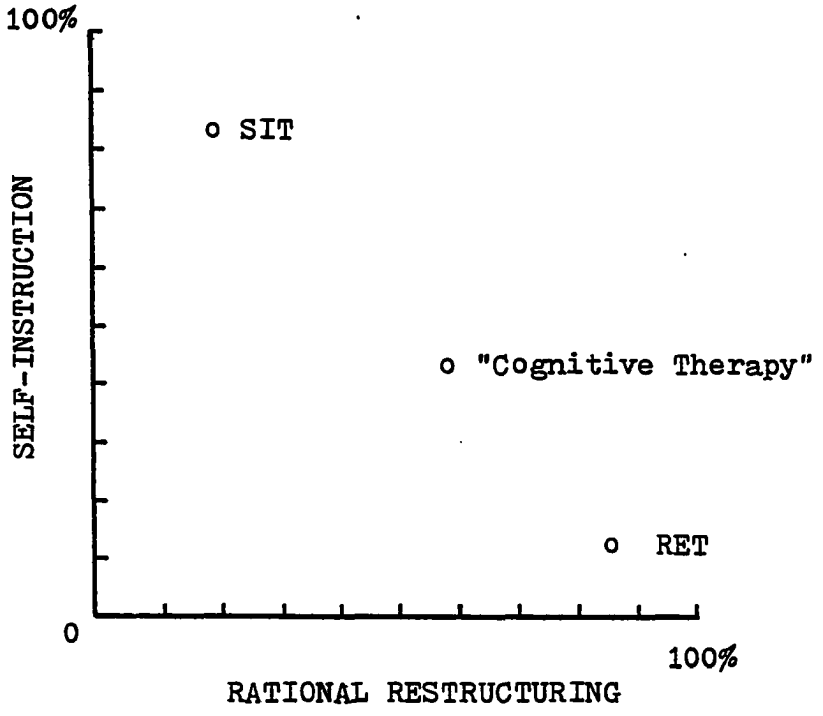


Figure 1. A figurative portrayal of differences among cognitive restructuring therapies.

restructuring techniques were to be included, provided that some element of self-statement modification was incorporated.

Method of Evaluating Outcome

In most direct terms, the goal of the present integration of research was to determine the parameters which moderate the effectiveness of cognitive restructuring techniques. The intent is most adequately expressed in the question posed by Paul:

> What treatment, by whom, is most effective for this individual, with what specific problem, under what set of circumstances? (1969, p. 162)

The present section deals with a brief description of the meta-analytic techniques for evaluating the variables which are suspected to influence outcome.

Generally, when one performs the aforementioned operation of converting outcomes of the studies to ESs, the discussion of meta-analytic procedures is largely reduced to a discussion of the statistics commonly employed in primary analyses. That is, effect size becomes the datum, and most common statistical procedures become applicable. For example, Pearson correlation coefficients can be calculated to describe the relationship between outcome (ES) and any set of corresponding observations, such as length of therapy.

As previously suggested, one qualifying distinction of meta-analysis is that statistical significance levels in the

usual sense are often inappropriate. The studies in a meta-analysis usually represent all available investigations within a selected area, and thus do not comprise a "random sample" from a "population" in the sense upon which most statistical indices are based. The studies may reflect what Rosenthal (1979) has termed "the file drawer problem," referring to the bias typically operative in any body of research because of the tendency for only "significant" results to be published. Finally, ESs may not be strictly independent observations, a concern which arises when studies contribute unequal numbers of outcomes to the meta-analysis.

The problems of inference in meta-analysis can be dealt with in three general ways: (a) reliance on descriptive information; (b) use of standard inferential procedures under conditions of data structure which satisfy the necessary assumptions of the procedures; (c) employment of inferential procedures which take the meta-analytic context into consideration.

Descriptive information, which is often underrated in primary analyses in favor of significance testing, takes on an even broader value in meta-analysis. Since the entire set of relevant studies to date are included in the analysis, many of the appropriate research questions are likely to address differences that exist <u>within</u> the sample. Levels of significance would not necessarily be relevant under such circumstances. Furthermore, meta-analysis sample sizes are often so large that

significance testing reduces to "an empty pro form ritual" (Glass & Smith, 1979, p. 8).

The relationships inherent in the data can be effectively addressed in a number of ways. A productive device in earlier analyses has been the expression of effects in terms of the relative shift in position implied--with reference to the distribution of scores--attributable to the treatment in question. For example, Smith and Glass (1977) determined that the average psychotherapy treated subject was better off than 75% of the untreated subjects. Similarly, Kulik et al. (1979) found that the Personalized System of Instruction raised the performance of the typical student from the 50th to the 70th percentile. It also seems clear, in light of the aggregative sample sizes involved, that the size of such effects would not be usefully supplemented by attachment of a "significance level".

Aside from categorical distinctions and subsequent comparisons of percentiles or mean differences, correlational techniques are also of direct value. Here, when appropriate, significance levels may be helpful in gauging the size of effects which might be expected by chance alone, if small effects or large numbers of variables are to be considered. Thus Smith and Glass (1977) concluded that variables such as client I.Q. and similarity of clients and therapists related significantly ($P < .05$) to psychotherapy outcome. In a similar vein, analysis of variance or, with more stringent restrictions,

multivariate analysis of variance may also be useful in determining the proportions of variance accounted for by certain variables or interactions of variables.

Various prospects for the application of multiple regression techniques are of particular relevance to the present research questions. In terms of outcome, multiple regression techniques can address the interrelated effects of various predictors, the hierarchy of importance of predictors, optimal combinations of predictors, the relationship of one or more variables to outcome controlling for (partialing out) the effect of one or more other variables, etc. For the present data, examples are numerous, such as (a) specification of an optimal set of therapist qualities, (b) specification of the methodological flaws which most strongly bias outcome, or (c) evaluating effectiveness controlling for design or analysis differences. Clearly Paul's (1969) observations call for several considerations of simultaneous or hierarchical sets of predictors.

Finally, Rosenthal (1976, 1978) has advocated a group of procedures allowing statements of confidence levels appropriate for the meta-analysis context. In essence, the procedures dictate several ways of mathematically combining the probability statements of the independent studies to form an overall probability statement associated with the effect in question.

Rosenthal (1978) discusses the relative merits and limitations of each approach, and argues for consideration of the magnitude of effect and confidence intervals around the

mean effect size as well as the overall probability statement. The latter can take the form of calculating the number of studies it would take to reverse one's conclusion if all new findings were in disagreement. This figure, qualified by ES and sample size, can lend credence to conclusions drawn,and addresses the viability of such alternative hypotheses as the "file drawer problem". In keeping with Rosenthal's recommendations, and the technology advanced by Smith et al. (1980), the present investigation addresses therapeutic outcome and its determinants both at the descriptive level and, where appropriate, at an inferential level.

METHOD

Selection of studies. The present intention was to perform a meta-analysis of all controlled, clinical investigations of the outcome of cognitive restructuring procedures which incorporate self-statement modification. The minimum requirements for inclusion of a study in the meta-analysis directly follow:

1. At least one treatment group must explicitly represent the application of a SIT variety of cognitive restructuring.
2. The cognitive treatment must be compared to at least one control group.
3. The treatment must be applied to the alleviation of some clinically relevant "problem". Analogue studies, applications as a learning tool in the classroom, etc., were excluded.

Consistent with the tradition of meta-analysis (e.g., Smith & Glass, 1977), other qualifications of the studies, such as the adequacy of design and analysis, the number of supplemental techniques, or the degree or orthodoxy represented in the application of the particular technique, were not used as criteria for inclusion, but rather as variables to be related to outcome. As Glass (1976) has noted, a poorly designed or analyzed study is not necessarily an invalid one: The relationship between methodological adequacy and outcome

is an empirical question. Similarly, an intervention which purports to follow a particular school of thought but is innovative or otherwise varies conceptually or procedurally within limits may be of no less importance in its findings. Meta-analysis allows such variations to be coded and treated, effectively, as empirical questions.

Studies were identified from several sources: (a) available literary reviews, (b) computerized literature searches, and (c) reference lists from studies already identified. Computer searches of Psychological Abstracts, Dissertation Abstracts, and ERIC were conducted with all appropriate variations on the keywords "cognitive restructuring, self-statements, cognitive behavior therapy/modification, stress inoculation, personal science, problem solving therapy/training, anxiety management," and "self-instructional therapy/training".

Over 1000 references were initially generated. Upon review of abstracts about 75% were found to be non-empirical, uncontrolled, or otherwise inappropriate for present purposes. A closer reading of 276 studies produced a sample of 113 investigations meeting all of the criteria for inclusion. Exclusions at this stage generally reflected lack of explicit self-statement modification, use of RET, or lack of a clinical "problem". Ten additional studies were sought for review but not obtained, primarily unpublished or extended reports for which requests to the authors produced no reply.

The 113 studies meta-analyzed (see starred references)

were principally oriented to Meichenbaum's (1977) SIT (73%). The remaining studies were of various orientations, including Beck's (1976) Cognitive Therapy. All, of course, made explicit reference to self-statement modification. The source of studies was quite closely balanced between published articles (47%) and dissertations (42%), with the remainder originating from masters theses and unpublished reports. Literature searches ended in November of 1980. The studies selected ranged in dates from 1968 to 1980, with over half from 1977 or later.

Characteristics of studies coded. Glass and Smith (1979) have noted that (a) a crucial factor in conducting a meta-analysis is the selection of variables which will be coded from the individual studies and used as independent variables or covariates for effect size, and (b) the most effective way to arrive at the set of variables is through familiarity with the substantive area, consultation with others familiar with the literature, and a careful initial reading of all studies to be included.

Through the process of reviewing literature in the areas of psychotherapy research, cognitive behavior modification and meta-analysis, a large number of suitable variables were identified. Consultation with senior colleagues aided refinement of the choice, format, and prioritizing of variables. Initial reading and pilot codings of selected studies lead to still further refinement. Coding schemes evolved and

unanticipated variables common across studies were added. Some variables which proved to be rarely observable were discarded. Only variables for which reliable, objective coding criteria could be developed were retained.

The coding sheets used, developed after initial piloting, are presented in Appendix A. The final set of variables coded for each study, including refinements, are portrayed categorically in Table 2. Coding criteria, where not readily apparent, are detailed in Appendix B.

The variables collected are similar to those in previous meta-analyses of psychotherapy outcomes (Miller, 1979; Smith & Glass, 1977), which also attempted to encode a wide range of characteristics which might relate to ES. Many of the present coding schemes and formats are also similar. There are major features of the present matrix of variables, however, which are unique to the present analyses.

First, as the present analysis focuses on a particular type of therapy, the components and features of each application were elaborated with much greater specificity. Consequently, for each cognitive treatment group a multiple classification scheme identifies any combination of at least 19 distinct therapeutic components of the technique applied. The purported theoretical orientation is retrieved, as well as the orthodoxy in its operationalization.

Secondly, reliability of variables is addressed differently in the present analyses. Coding large numbers of characteristics

Table 2

Characteristics of Studies Coded

I. Variables Coded Once for Each Study

General Identifying Information

1. Study identification numbers
2. Document date
3. Document source
   a. Journal
   b. Dissertation
   c. Thesis; unpublished report
4. Confidence in coding the study
   a. All necessary information was explicit
   b. Necessitated estimation conventions or missing data

Subject Characteristics

5. Age
6. Percent female
7. Percent minority
8. Intelligence (IQ)
   a. Mentally retarded
   b. Below average
   c. Average
   d. Above Average
9. Confidence of IQ classification
   a. Estimated from population characteristics
   b. Estimated from stated subject characteristics

Table 2 (continued)

c. Stated explicitly in the study

10. Socioeconomic status (SES)
    a. Low
    b. Average
    c. High
11. Confidence of SES classification
    a. Estimated from population characteristics
    b. Estimated from stated subject characteristics
    c. Stated explicitly in the study
12. Students?
13. Primary problem
    a. Neurotic
    b. Simple phobia
    c. Complex phobia
    d. Affect/depression
    e. Character disorder
    f. Delinquent/criminal
    g. Somatic
    h. Impulsive/hyperactive
    i. Handicapped
    j. Unassertiveness
    k. Discrete behavior
    l. Psychosis
    m. Antisocial behavior
    n. Stress/anxiety
14. Diagnosis method
    a. Self
    b. Psychometrics/extreme score
    c. Clinical
15. Type of phobia (if phobic)

Table 2 (continued)

a. Speech

b. Tests

c. Animal

d. Other

16. Length of hospitalization (months)

17. Method of soliciting subjects

a. Autonomously requested treatment

b. Response to ad

c. Referred

d. Experimenter solicited

e. Committed

Experimenter and Treatment Characteristics

18. Orientation of training

a. Psychology

b. Education or other

19. Treatment orientation

a. Meichenbaum

b. Beck

c. Other

20. Treatment similarity to Meichenbaum's

a. Meichenbaum was an author

b. Meichenbaum's work was the stated orientation

c. Multiple and/or other orientations

Table 2 (continued)

21. Number of therapists
Design Characteristics
22. Blinding
a. Experimenter(s) did the therapy
b. Experimenter(s) did not do therapy
c. Outcome assessed by blind collaborators
23. Subject group assignment
a. Random
b. Stratified random
c. Matching/equating
d. Convenience
24. Therapist assignment
a. Single therapist
b. Nonrandom, noncrossed
c. Random
d. Fully crossed
e. Matching
25. Counterdemand utilized?
26. Internal validity (see Appendix B)
27. Overall treatment mortality
28. Overall control mortality
29. All groups of same modality?
30. Equivalence assurance method

Table 2 (continued)

a. None, or placebo of measured inequivalence

b. Placebo without measurement of equivalence

c. Placebo of measured equivalence

31. Number of comparison groups

32. Number of cognitive treatment groups

33. Number of placebo groups

34. Number of alternate treatment groups

35. Number of group comparison combinations

36. Type of alternate treatment represented

a. Excluded cognitive techniques

b. RET

c. Traditional education

d. General counseling

e. Psychodynamic

f. Client centered

g. Gestalt

h. Relaxation

i. Systematic desensitization

j. Other

37. Treatment setting

a. School

b. Hospital

c. Clinic

d. Residential facility

e. Prison

f. College laboratory

Table 2 (continued)

II. Variables Coded for Each Outcome

General Identifying Information

38. Outcome identification numbers
39. Number of outcomes for this comparison
40. "Treatment" for this comparison
    a. Cognitive
    b. Alternate treatment
    c. Placebo
41. Type of alternate treatment (see #36)
42. "Control" for this comparison
    a. Placebo
    b. Intact group
    c. Wait list
    d. No treatment
43. Are multiple studies reported?

Subject Characteristics

44. Per cent female: treatment
45. Per cent female: control
46. Treatment sample size
47. Control sample size

Experimenter Characteristics

48. Allegiance to treatment
    a. Positive

Table 2 (continued)

b. Negative

c. Unknown

49. Treatment therapist experience (years)

50. Confidence of treatment therapist experience

a. Estimated

b. Stated

51. Control therapist experience (years)

52. Confidence of control therapist experience

a. Estimated

b. Stated

c. No control therapy provided

53. Treatment therapist sex

54. Control therapist sex

55. Treatment therapist/client similarity

a. Low
b. Medium
c. High
d. Very high

56. Control therapist/client similarity

a. Low
b. Medium
c. High
d. Very high

Treatment Characteristics

57. Treatment modality

a. Individual

b. Group

Table 2 (continued)

c. Mixed

58. Control Modality

a. Individual

b. Group

c. Mixed

59. Treatment number of sessions

60. Treatment duration (hours)

61. Treatment span (weeks)

62. Control number of sessions

63. Control duration (hours)

64. Control span (weeks)

65. Cognitive treatment components present (all that apply)

a. SIT?

1. Present

2. Primary (nominal) mechanism

b. Cognitive restructuring?

1. Present

2. Primary (nominal) mechanism

c. Modeling?

1. Present

2. Primary (nominal) mechanism

d. Homework; incompatible responses?

e. Behavioral rehearsal?

f. Role playing?

Table 2 (continued)

g. Imagery?
h. Covert self-reinforcement?
i. RET?
j. Stress inoculation training?
k. Self-instruction content manipulation?
l. Other cognitive techniques?
m. Social skills training?
n. Problem solving training?
o. Relaxation training?
p. Systematic desensitization?
q. Response cost?
r. Other behavioral?
s. Other?

Measurement Characteristics

66. Outcome type

a. Fear/anxiety (state)
b. Self-concept
c. Somatic/physiological
d. Sociopathic behavior
e. Discrete behaviors
f. Psychotic behavior
g. Public speaking
h. Depression
i. Impulsivity
j. Adjustment indicator
k. Personality traits
l. Cognitive task
m. Social behavior
n. Trait anxiety
o. Vocational
p. Assertiveness

Table 2 (continued)

q. Other state affect	r. Cognitions

67. Measurement sphere
    a. Behavioral
    b. Self report/cognitive
    c. Physiological
68. Reactivity of measurement

a. Low	c. High
b. Medium	d. Very high

69. Number of days after treatment measured
70. Reliability (if stated)
71. Reliability Type
    a. Inter-rater
    b. Split-half, test-retest, etc.

Outcome Characteristics

72. Treatment group mortality
73. Control group mortality
74. Equivalence
    a. Low
        1. Nonplacebo control
        2. Placebo of measured inequivalence
    b. Placebo
    c. Placebo of measured equivalence
75. ES calculation method

Table 2 (continued)

76. Means provided
    a. Post-treatment only
    b. Pre-post
    c. Adjusted, residual gains, etc.
    d. None
77. Report significance of effect
78. Treatment pre mean
79. Treatment pre standard deviation
80. Treatment post mean
81. Treatment post standard deviation
82. Control pre mean
83. Control pre standard deviation
84. Control post mean
85. Control post standard deviation
86. T
87. F
88. Mean square within
89. Weighting (if ES represents multiple outcomes)
90. Effect size
91. Significance of other effects addressed
    a. Subject sex
    b. Treatment credibility
    c. Age

Table 2 (continued)

d. Expectancy of improvement
e. Therapist
f. Other
92. Analysis of covariance used?
93. Average effect for this comparison
94. Average effect for this study

of studies confronts a pervasive lack of uniformity in what is reported and how it is reported. Even within the same journal comparable data are rarely found. While in one study IQ of subjects may be stated, in another it may go unmentioned. As in previous meta-analyses, coding schemes were developed which can often estimate parameters not directly given (see Appendix B). Thus, "above average" IQ may be inferred from the fact that all subjects in a study were college students. Still, McGaw and White (1980) have cautioned against letting such estimation conventions become too subjective or too far removed from the information actually reported. Moreover, even if reliable, the different levels of inference required may alter measurement properties of the variable or threaten its validity as an indicator of the property sought.

Presently, variables were retained only if an objective coding scheme could be devised using information directly stated in the studies under review. Global or subjective judgments beyond stated information were avoided. Reliability of coding should be limited only by accuracy of reading. Additionally, as suggested by McGaw and White (1980), the "confidence level" in coding variables with multiple levels of estimation was recorded (see Table 2). When appropriate, variance contributed by levels of inference in coding such variables was analyzed and statistically partialled out.

The execution of a meta-analysis hinges upon converting all outcomes to a standard unit of measurement, ES, and

relating ES to properties of the outcomes they represent. It is difficult to anticipate with precision the specific course of statistical analyses which will ensue once the data are coded and organized. Much of the analysis is sequential, and the nature of findings at any one step can greatly influence the pattern of subsequent analyses and the choice of statistical procedures which might be employed. It is clear at the outset of the present meta-analysis that three levels of information are desirable relevant to the principal research questions already posed:

1. Descriptive analyses: Summary statistics which will clarify the typical ES across all studies and within meaningful subroups (e.g., placebo controls versus no treatment controls, cognitive restructuring versus alternate treatments, etc.). Finer breakdowns of mean effects, inspection of scatterplots, correlational analyses and other procedures as indicated may help to delineate the sources of ES variability.
2. Regression analyses: Groups of related variables, such as therapist or subject characteristics, may be simultaneously regressed on ES to determine their constellations which optimally predicted outcome.
3. Inferential analyses: Where appropriate, probability estimates may be attached to the sources of differential effectiveness observed.

Placebo groups. The present meta-analysis differs from

precedents in several features of methodology. Foremost, placebo groups were treated in a much different manner. Smith and Glass (1977) analyzed placebo groups as a separate category of treatment, contrasting them against no treatment controls in the same manner as other treatments. Their results suggested a smaller effect on the average for placebo groups ($\overline{ES}$ = .56) than for other, conventional forms of psychotherapy ($\overline{ES}$ = .68).

While Smith and Glass were obviously sensitive to the importance of placebo groups, their analyses were incomplete, ending with calculation of the average placebo effect. They concluded that placebos were less effective; further analyses and discussion dwelt on the contrast between psychotherapy and no treatment controls.

The need for placebo groups gained early attention in Rosenthal and Frank's call for research paradigms including a "therapy in which clients have equal faith but which would not be expected by the theory of therapy being studied to produre the same effects" (1958, p. 471).

The technology of placebo group designs has since steadily evolved, although many of the critical developments have emerged after studies meta-analyzed by Smith and Glass (1977). The demand for adequate controls for expectancy effects, demand charateristics or other nonspecific effects persisted and grew, prompting innovative efforts to devise placebo groups more equivalent in credibility, attentional

value, etc. (e.g., Paul, 1966). Researchers have questioned the success of efforts in this regard. In reviewing systematic desensitization studies, generally considered to be a rigorous body of therapy research, Kazdin and Wilcoxon (1976) found that few studies presented any actual evidence of equivalence of comparison groups on expectancy or credibility factors. Consideration of only studies with stringent controls of assured equivalence eliminated most of the array of data in support of treatment efficacy. O'Leary and Borkovec (1978) have questioned the notion of "nonspecific effects" in the design of psychotherapy placebos, arguing that the controlled for factors pose a complex mass of active "therapeutic" ingredients, albeit difficult to assemble or maintain and poorly understood. The difficulties in placebo design and the ethical concerns in subjecting clients to placebo treatment have lent support to exploration of other types of comparison groups, such as a standard "minimally effective" treatment group against which new treatments can be contrasted (O'Leary & Borkovec, 1978).

Meta-analysis cannot bypass the problem of appropriate comparison groups (Dush, Tout, & Gabriel, 1980). In the Smith and Glass (1977) meta-analysis (a) not all studies had placebos, (b) no comparisons were made directly between placebos and psychotherapies, and (c) the characteristics accounting for the variability of placebo effects, with reference to either no treatment controls or psychotherapy, were not assessed. Given the implications of recent literature, a disturbing

possibility emerges: Effect sizes reported, while accurate in magnitude, may not be reflecting the true effect of interest. The precedent of Kazdin and Wilcoxon (1976) suggests that the Smith and Glass finding of an impressive (.68) overall mean psychotherapy effect may actually have little bearing on each technique's supposed active ingredients. One may simply or primarily have a .68 effect attributable to a concentration of positive credibility, attention, expectation of change, contact, and other nonspecific factors in the therapy groups. Precedent argues against dismissing these concerns.

The implication is not that no evidence for the efficacy of psychotherapy exists. As Glass (1976) notes, the relationship between characteristics of studies and their outcomes is an empirical question. Moreover, the technology of control groups implies that the definitive test of psychotherapy is a comparison of treatment groups against adequate placebos. Such an analysis has yet to appear.

In the Smith and Glass (1977) meta-analysis, the most instructive set of comparisons was not made. No effect sizes were calculated by treating placebos as "controls" in direct contrast to therapy. The overall difference in ES can be estimated, however. With common reference to non-placebo controls, "therapy's" $\overline{ES}$ of .68 compares with "placebo's" $\overline{ES}$ of .56. If we generously assume that placebos were equivalent to treatments on all appropriate dimensions of importance

we are left with little advantage of using therapy versus placebo. Referencing a normal distribution, while "treated" subjects were better off than 75% of controls, placebo subjects were better off than 71%! Roughly speaking, therapy appears to move the typical placebo subject only from the 50th to the 55th percentile, an effect of little compensation for the expense and professional investment incurred. It is likely that placebos were generally lower on such factors as expectancy, credibility, attention, etc. If properly assessed these might easily consume the remaining advantage of psychotherapy.

The present investigation was designed to extensively address these concerns. Many of the studies under review utilized placebo groups. Generally the definition and design of placebos followed the aforementioned notions of Rosenthal and Frank (1958). Such control groups were separately coded and analyzed in the present analyses. Procedures, composition, and other characteristics of the placebos were fully coded and statistically analyzed, including the degree of equivalence between placebo and therapy documented. Comparisons were made directly to both therapy groups and, where present, no treatment controls of the corresponding study, with all groups coded on comparable variables. The interrelationship of placebo effects and treatment effects was addressed in the context of all analyses, such as the influence of therapy type, length of treatment, reactivity of measurement, etc. In effect, the present design represents two parallel meta-

analyses: Psychotherapy and placebos versus no treatment controls, and psychotherapy versus placebos.

Covariates of effect size. A complexity of meta-analysis warranting more purposeful attention is treatment of ES covariates (Dush, et al., 1980). Each datum of the meta-analysis, as an ES of its corresponding individual ("unit") study, is an end product of all factors influencing the findings of that study. In practice these factors always encompass both "true" effects of interest and any of the biases injected by inadequacies of the study's design, measurement, implementation, etc. Meta-analysis is thus necessarily a multi-level, "meta-" methodology, needing to address both (a) the soundness of its aggregative procedures and (b) the integrity of the procedures which originally produce the data, as well as (c) the interaction of the two.

The same complexities are confronted in the traditional, narrative integration of research, but according to meta-analysis proponents, with inherently insufficient rigor, objectivity, and parsimony so as to negate the utility of the literary review (Glass, 1976). In contrast, meta-analysis promises a staightforward empirical treatment of many of the sources of error reflected in findings aggregated across studies.

While the resource for treatment of ES covariates in meta-analysis is of promising utility, it can easily be taken too lightly. The inevitable complexity of multivariate relation-

ships in aggregated data, and the multiple levels of analysis implied, pose difficulty in arriving at valid interpretations of covariance in the meta-analytic context. Variables coded to address the methodological flaws of unit investigations are of particular concern as they pose the greatest potential threat to the validity of the meta-analysis. However, the issues raised apply in large measure to any type of variable coded.

In terms of outcome research, this reduces to the traditional problem of "suppressor relationships." That is, the apparent correlation of variables X and Y (r_{xy}) may not be strictly interpretable, meaningfully, if a third variable Z exists such that $r_{xz} \neq 0$ and $r_{yz} \neq 0$. If we "partial out" the influence of Z or X and Y, the value of r_{xy} may change dramatically within limits determined by the magnitude of r_{xy}, r_{xz}, and r_{yz}.

The tenets of correlational analysis need not be belabored, but an example of the implications may be instructive. In the Smith and Glass meta-analysis, methodological characteristics of therapy outcome studies such as blinding, sampling versus matching, etc., were compiled into a composite "internal validity" variable, which was in turn found to correlate near-zero with the 833 aggregated ESs. It is tempting to conclude that the internal validity of unit investigations thus had no overall bearing on therapy outcome results, and rule it out from further consideration. Indeed

once cited, the internal validity variable received no further comment (Smith and Glass, 1977).

In actuality, the near-zero correlation gives little reassurance if scrutinized. Source of the study (published versus unpublished), experience of investigator, or any of a number of feasible intervening variables could mediate the influence of internal validity in intractible ways. For example, a zero correlation between internal validity ratings and ES could reflect a highly negative bias for published studies (perhaps selected by a tendency to publish significant results) and a highly positive bias among unpublished studies. In this case, the correlation between ES and the source of study could also be highly distorted if calculated without consideration of the influence of internal validity. To conclude that "overall, internal validity has no effect on outcome," would be a ludicrous distortion in this extreme case. Unfortunately, the assurances are not substantially more secure in a meta-analysis which does not endeavor to consider the influence of all possible mediators.

Realistically, there are always constraints to be contended with on time and the number of predictors and analyses any investigation and investigator can readily tolerate. Nonetheless, it is precarious to avert exhaustive multivariate consideration of prioritized covariates which have a broad potential for threat to the validity of findings. Moreover, temptation to implicitly or explicity "rule out" the

influence of a variable on ES, given an insubstantial zero order correlation, is generally without accountable foundation. Present analyses reflect the contention that variables such as internal validity, with potential for pre-emptive erosion of confidence in findings, should be at least initially considered as covariates in each phase and components of subsequent analyses (i.e., "partialing out" influence on the correlation between ES and other coded variables, routine inclusion as predictors in regression analyses, etc.).

Other design features. Several other design and analysis features unique to the present meta-analysis of outcome studies can be briefly summarized.

(a) Effect sizes were calculated only from observable data; Smith and Glass (1977) derived formulae to estimate ES from reported analysis of co-variance adjusted means and their variability. However, this requires the intercorrelation of outcome measures of outcomes and covariates, which is very rarely cited in studies. In such cases Smith and Glass arbitrarily estimated these inter-correlations. Rather than risk introducing systematic error in this regard, present analyses utilized the derivation formulae discussed by Glass (1978) with the most powerful data actually reported. Often, this reduced to an estimate of

ES from reported significance levels if only analysis of covariance results were reported.

(b) Only control group designs were analyzed. Smith and Glass estimated ESs for studies contrasting only two psychotherapies by normative reference to other controlled studies of similar psychotherapies. Again, this removes ES further from a straightforward standardization. Present analyses dealt only with studies employing at least one placebo or control.

(c) The present analysis did not rely solely on the Glass (1976) approach to meta-analysis. As previously detailed, Rosenthal's (1978) combined probability methods were applied as appropriate.

(d) Present analyses addressed the distribution characteristics of observed ESs (variance, skewness, kurtosis, etc.), which have been shown to bear on both the Glass (1976) and Rosenthal (1978) meta-analytic procedures (Tout & Dush, 1980).

(e) As in previous meta-analyses, many studies were found to have multiple measures of the same construct (e.g., state anxiety). In the present analysis ES's were calculated for all measures for all comparisons of all studies (as opposed to sampling, which has been done on occasion in earlier works). When multiple measures of a construct identical on <u>all</u>

coded study characteristics were encountered, these were combined and only their average was recorded. Subsequent analyses of mean ES accounted for this operation's influence by weighting the averaged entry by the number of separate original outcomes it represented. Thus, all outcomes contribute equally to any aggregation of mean effect. Influences on the variance of ESs were accounted for by using the weighting factor as a covariate in correlational analyses, statistically partialling out its influence where appropriate.

RESULTS

Findings from all Studies

The 113 studies coded represented 345 separate comparisons of "treatments" versus "controls", using placebos as both controls (relative to psychotherapy) and treatments (relative to no treatment controls when present in the same study). A total of 2335 treated subjects were represented in the 217 comparisons against no treatment; the typical comparison contrasted 10.76 treated subjects against a no treatment control group of 10.41. The 128 comparisons of psychotherapy versus placebo contrasted 1305 psychotherapy subjects (mean = 10.20) against an average placebo group of 9.73 subjects. Overall, 40% were female, 47% were of above average IQ, and 80% were students, with an average age of 17.8. The modal age (19) was typical of the heavily represented group of college students.

Comparisons reflected 1786 ESs against no treatment controls and 1070 ESs against placebos, for an average of 8 per comparison. A total of 331 outcomes were associated with placebo groups which were analyzed both as controls and, in separate analyses, as interventions contrasted against no treatment controls from the same studies. While analyses are kept separate, it bears noting that these placebos are analyzed twice.

Allowing for the dual consideration of some placebos, each study produced an average of 22.35 nonredundant ESs. This was

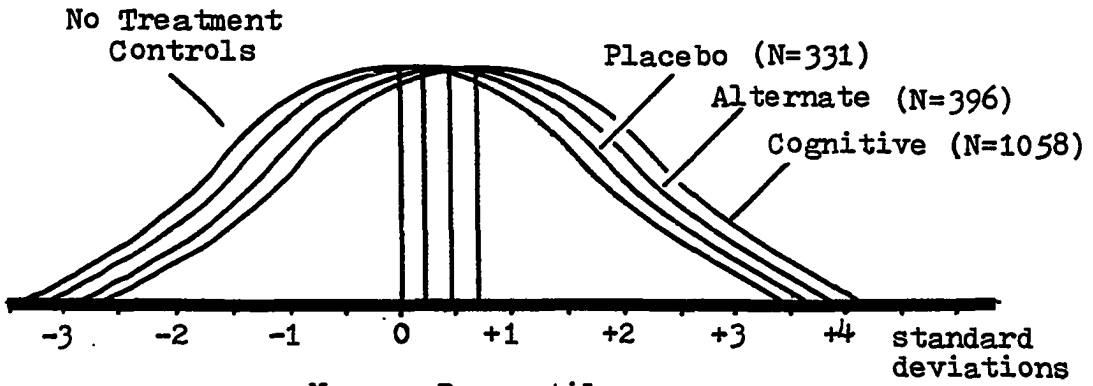
unanticipated, differing drastically from Smith and Glass' (1977) average of 2.22 ESs per study. The discrepancy may in part be due to more exhaustive ES coding presently. A more likely cause, however, is the contemporary trend of using more comparison groups per study and more outcomes per comparison.

Major aggregations. The average ES for all forms of treatment versus no treatment was .52. Referencing, for illustrative purposes, a normal distribution of untreated subjects, the typical treated subject was better off than 70% of those untreated. Stated differently, treatment, on the average, appeared to move a subject from the 50th to the 70th percentile.

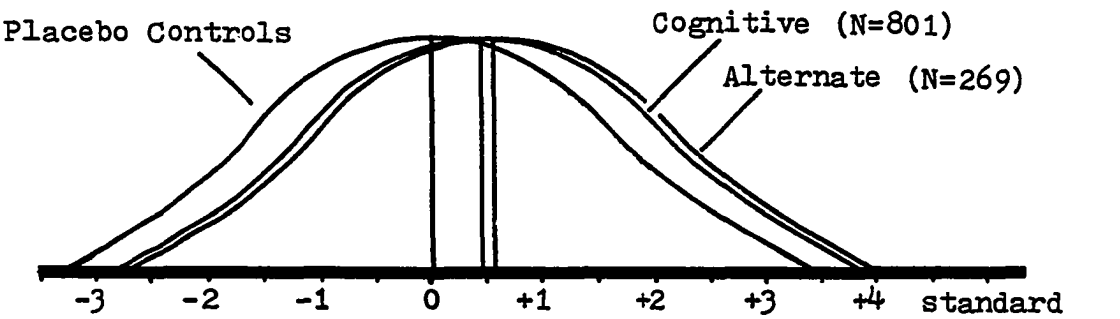
Differences in efficacy across major treatment categories were considerable. These are depicted in Figure 2 as shifts in ES distributions. Normal curves are again used for convenience. The observed distributions generally tended to be somewhat leptokurtic and positively skewed, although this may be attributable to the infrequency of strongly negative outcomes and a disproportionately high frequency of zero outcomes when ES could only be estimated from reported tests of significance. Normality may be a reasonable assumption for approximate descriptive representations of therapeutic gain.

Cognitive restructuring was the most effective treatment, with a mean ES of .66. This can be gauged as a shift from the 50th to the 75th percentile. Notably, this shift is identical to the average effect of all varieties of psychotherapy meta-analyzed by Smith and Glass (1977). The overall

Figure 2. Improvement of typical treated subjects relative to the distribution of controls.



	Mean Effect	Percentile Shift	Skewness	Kurtosis[a]
Cognitive:	.66	75th	2.59	25.16
Alternate:	.45	67th	.91	5.80
Placebo:	.17	57th	-.68	6.61



	Mean Effect	Percentile Shift	Skewness	Kurtosis[a]
Cognitive:	.43	67th	1.66	7.87
Alternate:	.47	68th	1.89	6.33

([a] Kurtosis formula assigns the normal curve a value of 3.00).

average effect of all alternate psychotherapies represented was .45, corresponding to a shift from the 50th to 67th percentile. Placebos produced a relatively small effect ($\overline{ES}$ = .16) and a shift only from the 50th to 56th percentile.

The comparable set of comparisons between psychotherapy and placebo controls is also illustrated in Figure 2. Compared to placebo treatment, psychotherapy produced an overall effect of .44, for a percentile shift from the 50th to the 67th, referencing the placebo control distribution. Cognitive restructuring produced a mean effect of .43. Alternative treatments, with a mean effect of .47, emerged with a one percentile advantage over cognitive restructuring. As apparent in Figure 2, the difference is so small as to be, for all purposes, negligible.

The loss in advantage of cognitive restructuring over alternate treatments when controls are placebos is further addressed in Table 3. Principal categories of frequently represented alternate treatments are distinguished for both sets of comparisons (placebo and no-treatment).

No subgroup of alternate treatment produced a mean effect larger than the mean effect of cognitive restructuring therapies contrasted against no treatment controls. In comparisons against placebos alternate treatments fared much better. Traditional (classroom or tutorial) education, representing primarily child therapy, produced a very impressive $\overline{ES}$ of .97

Table 3

Breakdown of Alternate Treatment Mean Effects[a]

Treatment Type	No Treatment Controls			Placebo Controls		
	Mean ES	Percentile Shift[b]	Number of ESs	Mean ES	Percentile Shift[b]	Number of ESs
Education	.27	61st	25	.97	83rd	22
Psychodynamic	.52	70th	16			-
Relaxation Training	-.03	49th	23	.05	52nd	16
Systematic Desensitization	.59	72nd	96	.44	67th	70
Other Behavioral	.55	71st	160	.49	69th	115
Other Cognitive	.20	58th	35	.41	66th	36

[a]The standard deviation for ESs across all groups is .66 for comparisons against no treatment controls and .39 against placebos.

[b]Placement of the typical treated subject with reference to control subjects.

against placebo controls, where few large effects have yet emerged. Quite the opposite finding was evident in contrasts against no treatment controls. Education methods fared poorly in this case ($\overline{ES}$ = .27) while other alternate therapies, on the average, were more effective. Factors which may bear on this large, counterintuitive discrepancy will be examined in later sections.

In comparison against placebos, cognitive restructuring, other cognitive techniques, systematic desensitization, and other behavioral techniques were within a three percentile range, with little variance being accounted for by these categories. It is noteworthy that relaxation training, though modestly represented, consistently produced a near-zero mean effect. Without the linkage of relaxation to some form of anxiety stimulus hierarchy (the minimum requirement for categorization as systematic desensitization), little effect was realized.

Matched comparisons. It is clear that not all studies used alternate treatments. Consequently, the effects previously cited correspond most directly to the subset of ESs for cognitive restructuring which came from studies which include alternate treatment. Such comparisons of cognitive restructuring against no treatment controls produced 487 ESs, about half of all cognitive restructuring effects. The mean effect for this subset was .78 (a shift to the 78th percentile), in contrast to the .66 $\overline{ES}$ cited for all cognitive restructuring outcomes.

Comparisons of cognitive restructuring against placebos from studies including alternate treatment produced 329 ESs, with $\overline{ES}$ equal to .39 (65th percentile). This entails a slightly lower effect than noted for all 801 cognitive restructuring outcomes contrasted against placebos ($\overline{ES}$ = .43).

The overall implication is straightforward. Examining the effects from studies which incorporate both cognitive restructuring and alternate treatment magnifies previously cited patterns. For comparisons against no treatment controls, cognitive restructuring evidences a larger advantage over all forms of alternate treatment. In comparisons against placebos, cognitive restructuring lost ground, moving to a position of lower average effectiveness than three categories of alternate treatment. The differences here remain modest, however. At a 65th percentile shift, cognitive restructuring remains within four percentiles of the leading group of behavioral techniques.

Control groups. Additional breakdowns according to control group type are of interest. Within comparisons against no treatment controls, two types of control groups are encompassed: (a) Those where subjects were assessed within the context of being on a waiting list for treatment at a later date and (b) those of assessment only, with no promise of future treatment. About twice the number of ESs (N_{ES}) were associated with waiting list controls (1135) as opposed to assessment only controls (648).

Differential effectiveness between the two types of

no treatment controls would suggest differential rates of change from pre-test to post-test, perhaps attributable to the promise of future therapy. The mean ES of all treatments against assessment only controls was .38; against waiting list controls, the overall average ES was .60. A substantial difference thus emerges, but in the unexpected direction. Treated subjects improve more relative to waiting list controls than relative to subjects only assessed!

There is no adequate way to eliminate the possibility that other characteristics of studies explain this paradoxical effect. Unfortunately, virtually no data exist from studies using both waiting list and assessment only controls. A sample of studies cannot be selected that would be matched on other factors. It is difficult to conclude that with nearly 1800 observations at hand, the finding is the product of spurious, systematic differences between investigations using one type of control over the other. One such explanation presents itself, but is difficult to evaluate. It is possible that experimenters treating more severely troubled subjects, for ethical reasons, are more inclined to use waiting list controls. Matched or equated comparisons against both control types are needed to evaluate this or other explanations.

Exploration of Differential Effectiveness

Characteristics of studies considered relevant as potential

determinants of outcomes were analyzed in several ways. Categorical variables were evaluated by inspection of mean effects across each of their categories; continuous variables were correlated with outcome, partialling out the influence of weighting and validity indicators; meaningfully related groups of continuous variables were addressed in step-wise multiple regressions; scatterplots were inspected where appropriate for curvilinear relationships; dummy coded regressions, covarying for weighting and validity indicators, were used to assess the relationship of various categorical variables to ES. Simultaneous consideration of all variables in a multivariate fashion was impractical. Where feasible analyses did attend to the moderating influence of variables found salient in earlier stages of analysis.

Age. Age of subjects correlated near-zero with ES across all outcomes (r_{xy} = .04). Partialling out the influence of ES weighting further reduced the correlation ($r_{xy.w}$ = .03). Yet, conceptually, differences between adults and children are of considerable interest. The process of therapy would be expected to differ across these two age groups, as well as the nature of problems, outcome measures, and other characteristics. Moreover, variation across treatment and control types, fluctuation across some areas of the age range, and some appearance of curvilinearity were suspect of masking the influence of age.

These speculations were examined in breakdowns of $\overline{ES}$ distinguishing children (under 18) from adults (18 and over).

Table 4 presents the results for both placebo and no treatment control sets of comparisons. Different ages exhibited substantial differences in treatment effectiveness throughout all comparisons. Outcomes were quite consistently more favorable for adults than children. The one exception was in comparisons between alternate treatment and placebo controls. The small number of ESs associated with alternate treatments for children produced a mean effect (.65) higher than alternate treatments for adults (.45) and in fact, the highest mean effect of all categories within placebo control comparisons.

Aside from the same exception, cognitive restructuring was more effective in each category than alternate or placebo treatments. Its advantage with adults over alternate treatments grew markedly in contrast to no treatment. Once again, this advantage shrinks decisively when considering comparisons against placebo controls.

The results for children are discouraging. The mean effects are all lower than average with respect to mean effects in earlier meta-analyses. Placebos for children, though infrequently contrasted as treatments, produced virtually no effect on the average in comparison to no treatment. The larger placebo effect for adults in this respect poses an interesting point for consideration. Perhaps, in some ways, children are less gullible than adults.

Validity. Indicators of threat to internal and external validity were given early consideration. For reasons discussed,

Table 4

Age Differences for all Treatments[a]

Treatment ...Age	No Treatment Controls			Placebo Controls		
	Mean ES	Percentile Shift	Per cent of ESs	Mean ES	Percentile Shift	Per cent of ESs
Cognitive						
Children	.44	67th	16%	.27	61st	29%
Adults	.74	77th	43%	.53	70th	46%
Alternate						
Children	.27	61st	4%	.65	74th	3%
Adults	.49	69th	19%	.45	67th	21%
Placebo						
Children	.07	53rd	6%			-
Adults	.21	58th	12%			-

a relationship between validity and outcome, if extant, may bear on many aspects of subsequent analyses.

Several of the variables coded were designed to assess threats to internal or external validity of the unit study (Fiske, Hunt, Luborsky, Parloff, Reiser, & Tuma, 1970; Smith & Glass, 1977). Six of these are depicted in Table 5, with analyses of mean effects for all levels of each variable.

The "validity" variable is a bivariate composite developed by Smith and Glass (1977) capturing threats contributed by subject attrition. As detailed in Appendix B, its score is reduced for (a) high attrition, (b) unequal attrition across treatment and control groups, and (c) nonrandom subject assignment, in progressively undesirable combinations.

The results in Table 5 do not suggest a linear effect for the validity variable with no treatment controls. In comparisons against placebos, however, an interrelationship is evident. ESs from the lowest third of validity ratings are the largest in magnitude, nearly double the mean effect of higher validity studies. Interestingly, this threat appears in the more methodologically rigorous set of comparisons using controls equated on nonspecific factors. Perhaps in less rigorous designs some spurious sources of variance are masked or undetected.

The validity rating, as noted, further penalizes high and/or unequal attrition comparisons in an interactive manner if

Table 5

Validity Indicator Breakdowns

Variable Name	No Treatment Controls		Placebo Controls	
	Mean ES	Number of ESs	Mean ES	Number of ESs
Subject assignment				
Random	.48	721	.30	325
Stratified	.52	647	.57	431
Matched	.69	328	.41	226
Convenience	.19	89	.37	78
Therapist assignment				
Random	.55	363	.53	248
One therapist	.49	551	.44	325
Crossed	.50	442	.51	254
Other	.59	405	.27	225
Subject solicitation				
Self-referred	.69	21	.67	10
Response to ad	.81	532	.69	259
Referred	.24	261	.10	165
Solicited	.41	809	.42	491
Committed	.81	84	.71	70

Table 5 (continued)

"Validity"				
Low	.49	613	.61	317
Medium	.62	484	.34	285
High	.50	672	.39	453
Control equivalence				
Low			.13	47
Medium			.44	670
High			.48	345
Reactivity				
Low	.34	120	.40	58
Medium	.36	230	.44	152
High	.57	875	.37	544
Very High	.55	551	.56	302

random assignment is absent. The breakdown of subject assignment suggests that presence or absence of random assignment is indeed critical, even in isolation (see Table 5). For comparisons against either form of control group, random assignment was associated with markedly lower mean outcome than either matching or stratified random assignment. An exception is the small proportion of the outcomes from convenience samples with no treatment controls. However, if one collapses all categories of nonrandom assignment, the respective mean effects are .54 against no treatment controls and .50 against placebo controls. Once again more dramatic differences emerged from comparisons against placebos.

The various arrangements for assigning therapists to treatment conditions are depicted next in Table 5. Little meaningful difference of substance is apparent. Certainly the possibility that nonramdom assignment of therapists would result in favorably biased outcomes is not supported. Indeed, in the more rigorous comparisons against placebo controls, treatments by therapists assigned in a nonrandom manner and not crossed across conditions (not all treatments were provided by all therapists) had a lower mean effect.

Source of subjects is of interest as an external validity indicator to the extent that it relates to severity of disorders. Consequently, subjects institutionalized or referred by other caregivers for treatment would be expected to be more

relevant for generalization to clinical populations than those sought out by the experimenter for the purpose of research and induced to participate. Subjects presenting themselves for treatment might be expected to fall between the two extremes presently represented.

The pattern of mean outcomes (see Table 5) is interesting in light of expectations. The substantially represented category of "referred" subjects in both sets of comparisons reflects the lowest outcomes by a considerable margin. Institutionalized subjects, on the other hand, rank at the highest for both sets of comparisons. Subjects presenting themselves for treatment fared well comparatively, including those responding to ads for treatment. Experimenter solicited subjects, likely to be the least representative of clinical populations, responded less than the average of treated subjects. Generalizing from present findings to clinical applications would not appear to threaten undue biases of concern.

Comparisons against placebo controls varied in outcome in relation to the degree of equivalence assured between placebo and therapy on credibility, expectancy, and nonspecific factors. The direction of relationship is the opposite of that presenting concern to internal validity, with the most positive outcomes emerging from the best controlled comparisons.

Reactivity or "fakeability" of measurement, based on a Smith and Glass (1977) coding scheme detailed elsewhere (Appendix B) is also addressed in Table 5. Both sets of com-

parisons illustrate outcomes of some concern. The most reactive measures of outcome produced the largest effects. This difference is less pronounced with placebo controls, especially if we contrast the combined lower two categories ($\overline{ES}$ = .43) with the higher two ($\overline{ES}$ = .44). It is not surprising that outcomes with more stringent controls for comparison group equivalence would be less susceptable to measurement reactivity. Reactivity is one of the "nonspecifics" that is controlled in placebo group designs.

Aside from singular considerations of validity indicators a composite indicator is desirable for inclusion in later correlational analyses. A unit weighting scheme was adopted. All of the variables in Table 5 of ordinal scale were included in their original form (see Appendix A). Additional dichotomous variables were encoded to indicate presence or absence of (a) random assignment of therapists, (b) random assignment of subjects, (c) single blind measurement, and (d) experimenter solicited subjects. Similarly, dichotomous variables indicated the equivalence or nonequivalence of treatment and control on (a) length of treatment, (b) modality, and (c) therapist experience.

The overall correlation between this composite validity index (CVI) and ES, partialling out the ES weighting factor, was .01 for no controls and .05 for placebo comparisons. A very small overall relationship between the CVI and outcome is suggested. While reassuring, the previously noted possibility

of stronger, masked relationships within smaller subsets of data or suppression by other variables warrants continued consideration in subsequent correlational analyses.

Study characteristics. Additional characteristics of the unit investigations not directly related to validity were of interest. Many of these are considered categorically in the forthcoming sections; others may be located in various tables of correlational data presented.

Three variables which provide interesting information on the general context of studies are presented in Table 6: source of the report, training orientation of the principal experimenter, and the treatment setting.

Source of reports reveals a finding of concern (see Table 6). Published reports, from both sets of comparisons, are associated with more favorable outcomes. While of moderate magnitude, this discrepancy suggests caution in light of an apparent "file drawer" problem. Reliance on only published data would appear to give the clinician or researcher a distorted view of data bearing on therapy effectiveness. To the extent that the present investigation sought out unpublished, often obscure reports, some assurance against this distortion is provided in the meta-analysis.

Training of the senior experimenter was discernible for 88% of all outcomes. A large majority of these (86%) were identifiable as psychology trained, titled "clinical psychologist," or serving in a university department of

Table 6

Breakdown of Study Characteristics

Variable Name	No Treatment Controls		Placebo Controls	
	Mean ES	Number of ESs	Mean ES	Number of ESs
Source of report				
Journal	.65	767	.50	537
Dissertation	.40	786	.41	454
Other	.49	233	.19	72
Training				
Psychology	.49	1313	.50	856
Other	.55	255	.00	93
Location				
School	.20	277	.32	240
Residential	.65	101	.67	13
Hospital	.23	171	.35	165
Clinic	.72	50	.25	14
College facility	.61	1174	.55	619

psychology. Of the remaining 14%, most (93%) were similarly identified as of training from an education perspective, with the remainder principally from psychiatry. The differences in mean outcome proved to be provocative, even given consideration of the imbalance in number of ESs represented. While outcomes for nonpsychology experimenters show a slight edge in comparisons against no treatment controls, a large difference in the opposite direction emerged for placebo controls. In the latter instance a zero effect is observed for nonpsychology experimenters. Moreover, for the 83 ESs herein exclusively of education orientation, the average effect was -.01.

The differences observed are difficult to interpret. At least in part, these may be a reflection of other, not immediately apparent differences. Most suspect, for instance, would be subject age, since overall effects were lower for children and educationally oriented experimenters might be expected to be more involved with this age group. In actuality, educationally oriented experimenters did deal with children more frequently than the psychology oriented experimenters. The ratio between orientations was 2:1 with no treatment controls and over 3:1 with placebo controls. It remains unclear, however, why the large discrepancy emerges only in the more rigorous comparisons against placebo controls. Furthermore, for educationally oriented experimenters the mean effect against placebo controls was actually lower for adults ($\overline{ES}$ = -.05) than for children ($\overline{ES}$ = .00).

Location of treatment delineated considerable variability in mean effectiveness (see Table 6). Treatment within the school setting, once again reflecting the generally poorer outcomes for children, was relatively ineffective in both sets of comparisons. It is interesting that outcomes in the two most restrictive environments were consistently quite disparent. Outcomes in residential settings were high on the average, while typical outcome in the hospital environment was well below average. Treatment in college facilities, primarily reflecting treatment of college students, fared well throughout. It is interesting to note that treatment in mental health clinics, taken collectively, also fared well, drawing on a more heterogeneous array of subjects.

The publication year for each study was also considered. Simple correlation with ES was -.14 for no treatment controls and -.23 for placebo controls. Across all ESs this represents a correlation of -.18. Partialling out the influences of ES weighting and CVI had no bearing on this relationship. The temporal correlates of ES are illustrated graphically in Figure 4. The least-squares regression lines, of course, indicate a steeper decline in mean effectiveness for the more rigorous comparisons against placebos. It is also clear that declining trends have been much less consistent for comparisons against no treatment controls. The separate indications of means for cognitive restructuring alone confirm that the decline depicted by correlations and regressions is representative of

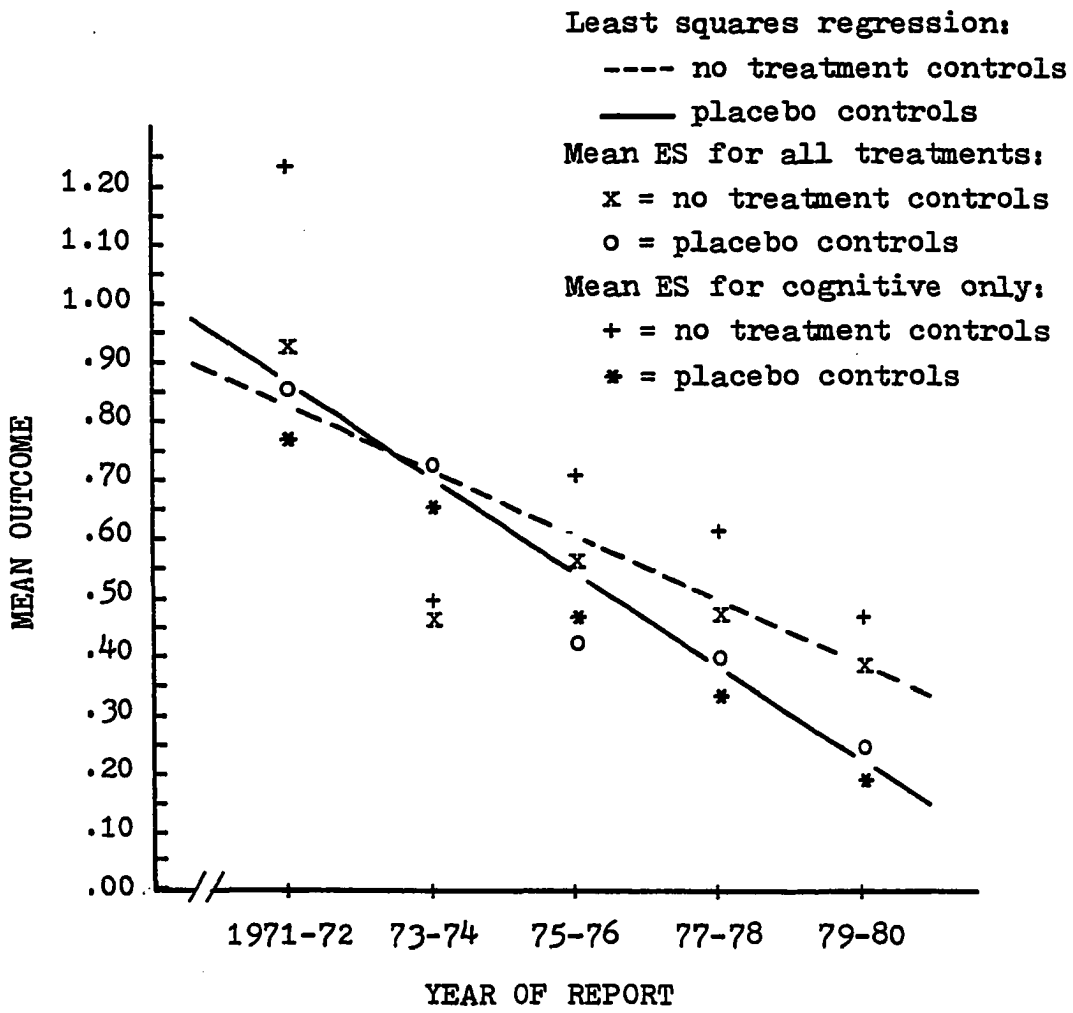


Figure 1. Observed and predicted outcomes by year of report.

cognitive restructuring outcomes as well.

Effectiveness across disorders. Numerous categories of disorders were treated within the scope of ESs represented. Among the most critical of concerns are the clinical implications of possible differential effects for different types of disorders. Table 7 presents an analysis of mean effectiveness for all major disorder groupings. It should be kept in mind that diagnostic labels were assigned according to the descriptions or labels used in the unit study. No effort was made at verification or systematic alteration. Table 7 reports analyses for small numbers of ESs in some subcategories. While these may be informative or suggestive, they are intended for cautious interpretation.

Categories with substantial numbers of ESs warrant especially close inspection. Child therapy was by far most frequently assessed in the treatment of impulsivity or hyperactivity (terms which were rarely distinguishable in the studies represented). It fared poorly in all instances except for the modest number of comparisons of alternate treatments against placebos. Taken collectively across controls, alternate treatment differed little from cognitive restructuring in overall effectiveness for this disorder.

The emerging general rule of relatively poor outcome in child treatment had some exceptions in analyses by disorder, generally in small groups of outcomes. Cognitive restructuring was well above the overall average in the

Table 7

Mean Effectiveness Across Disorders[a]

Disorder ...Treatment	No Treatment Controls				Placebo Controls			
	Children		Adults		Children		Adults	
	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}
Neurotic								
Cognitive	2.44	2	.53	27			.00	5
Alternate	.53	12						
Placebo								
Simple Phobia								
Cognitive	1.06	20	.67	413	.56	5	.46	210
Alternate	-.19	1	.49	161			.58	94
Placebo			.29	105				

Table 7 (continued)

Complex Phobia	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})
Cognitive			1.26	70	.00	2	.76	7
Alternate			.50	36			.83	7
Placebo								
Affect/Depression								
Cognitive			.77	46			.10	40
Alternate			.34	40			.13	30
Placebo			-.31	18				
Character Disorder								
Cognitive	1.02	9					1.04	2
Alternate	.55	9					.00	1
Placebo								
Delinquint/Criminal								
Cognitive	.70	27	.18	15	.20	6		
Alternate	.30	21			1.04	2		
Placebo	.49	7						

Table 7 (continued)

Somatic Disorders	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})
Cognitive			.69	24			.22	22
Alternate			.58	8			.83	4
Placebo								
Impulsive/ hyperactive								
Cognitive	.39	192			.31	253		
Alternate	.21	30			.62	32		
Placebo	.11	71						
Unassertiveness								
Cognitive	-.38	8	.80	126	-.37	8	.78	137
Alternate			.63	32			.47	54
Placebo	-.34	8	.17	56				
Psychosis								
Cognitive			.49	24			.68	40
Alternate			.23	27			.23	25
Placebo			.22	18				

Table 7 (continued)

Miscellaneous[b] behaviors	($\overline{ES}$)	(N_{ES})	($\overline{ES}$)	(N_{ES})	($\overline{ES}$)	(N_{ES})	($\overline{ES}$)	(N_{ES})
Cognitive	.09	32	.77	23	.10	35	.38	20
Alternate	.00	3	.86	16			.25	10
Placebo	.00	26	.26	10				

[a] $\overline{ES}$ = mean effect size; N_{ES} = number of effects represented.

[b] Includes various social, discrete, and adjustment behaviors.

treatment of neurotic, character, simple phobic, and delinquency disorders. The latter two were comparatively well represented in ES numbers among categories other than impulsivity. Treatment of unassertiveness and other circumscribed behaviors fared very poorly.

Two categories well represented for adults were treatment of phobias and unassertiveness. Both reveal mean effects exceeding overall averages for all treatment subgroups. Both categories of phobias show cognitive restructuring surpassing alternate treatments in comparisons against no treatment controls and a pattern reversal against placebos. Cognitive restructuring maintains a substantial advantage for unassertiveness across both sets of comparisons. In smaller numbers of ESs, similar patterns emerge for psychosis. The moderately represented categories for neurosis, somatic disorders, and affect/depression show an advantage for cognitive restructuring in comparisons against no treatment controls. The advantage is reversed for somatic disorders against placebos. Cognitive restructuring enjoyed above average effectiveness for other circumscribed behaviors against no treatment controls, although alternate treatments here produced an even more impressive mean effect of .81 for 16 ESs.

The coding scheme accounts for various subtypes of the frequently and effectively treated simple phobic disorders. These are illustrated in Table 8. Subject age was not distinguished since very few children were treated for phobic disorders.

Table 8

Treatment Effectiveness for Phobias

Phobia Type ...Treatment	No Treatment Controls		Placebo Controls	
	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}
Speech				
Cognitive	.64	150	.34	88
Alternate	.52	83	.56	52
Placebo	.23	51		
Test				
Cognitive	.66	239	.41	103
Alternate	.52	66	.23	15
Placebo	.38	50		
Animal				
Cognitive	1.06	21	1.81	13
Alternate	.47	9	.96	23
Placebo	.00	4		
Other				
Cognitive	1.31	79	.39	20
Alternate	.38	36	.53	11
Placebo				

The largest mean effects appear for the treatment of animal phobias (rodents, reptiles and insects). Cognitive restructuring here demonstrated a wide margin of advantage, though tested on a moderate number of ESs. Test anxiety, the adult disturbance highest in number of outcome measures, demonstrated a consistent advantage of cognitive restructuring. Mean outcomes were, nonetheless. roughly average for the treatment type. Cognitive restructuring outperformed alternate treatments for both remaining phobia types (especially in the residual "other" category) in the typical contrast against no treatment. In both cases mean outcome falls to below average in comparisons against placebos and is exceeded by alternate treatment.

Measurement characteristics. Differences observed across types of disorders are also likely to entail differences in the dimensions or types of therapeutic change being measured. Two kinds of information were collected to qualify ES measurement characteristics: (a) categorization of the content domain for each ES and (b) psychometric features of the measurement process.

Table 9 portrays the analysis of mean outcomes according to content domain measured. Here, most directly, the clinical nature of measured therapeutic change is revealed. The group of measures of affect encompassed very large numbers of outcomes. State anxiety was highest in the number of ESs. Over 300 outcomes produced a large (.76) cognitive restructuring effect

Table 9

Treatment Effectiveness by Measurement Domain

Domain ...Treatment	No Treatment Controls				Placebo Controls			
	Children		Adults		Children		Adults	
	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}
Fear/ State Anxiety								
Cognitive	1.06	18	.76	206			.47	117
Alternate			.33	77			.59	66
Placebo			.19	51				
Trait Anxiety								
Cognitive			.73	176			.38	94
Alternate			.60	65			.38	39
Placebo			.27	51				

Table 9 (continued)

Personality Traits (other)	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})
Cognitive	.15	1	.11	13			.00	1
Alternate	.92	1	.18	3			.00	1
Placebo			.00	1				
Adjustment Indicator								
Cognitive	.87	8	1.24	1			.65	1
Alternate	.10	8	.49	1			.65	1
Placebo			-.12	1				
Self Concept								
Cognitive	1.51	3	1.15	11	.00	1	.14	3
Alternate	.71	1	.67	9			.01	2
Placebo								
Somatic/Physiological								
Cognitive			.68	12			.15	10
Alternate			.47	4				
Placebo			.50	4				

Table 9(continued)

Psychotic Behavior	$(\overline{ES})$....(N_{ES})		$(\overline{ES})$....(N_{ES})		$(\overline{ES})$....(N_{ES})		$(\overline{ES})$....(N_{ES})	
Cognitive							1.64	10
Alternate							.28	6
Placebo								
Work/Vocational								
Cognitive			.95	7			.89	8
Alternate			-.27	4			.14	4
Placebo			-.26	4				
Public Speaking								
Cognitive	.45	1	.71	54			.34	22
Alternate	-.19	1	.40	28			.56	14
Placebo			.29	13				
Assertiveness								
Cognitive	.09	7	.85	107	-.01	7	.79	112
Alternate			.53	42			.57	44
Placebo	-.12	7	.23	53				

Table 9 (continued)

Depression	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})
Cognitive			1.07	23			.24	19
Alternate			.43	19			.05	12
Placebo			-.33	10				
Other Affect								
Cognitive	.26	11			.42	37	-.16	4
Alternate	.18	10			.09	2		
Placebo								
Achievement								
Cognitive	.34	96	.44	67	.20	120	.46	31
Alternate	.10	16	.36	20	.30	16	-.09	5
Placebo	.16	37	.54	11				
Cognitive Task								
Cognitive	.75	45	1.09	9	.68	55	.57	23
Alternate	.45	6	1.01	4	1.07	14	.25	14
Placebo	.21	16	.56	5				

Table 9 (continued)

Social Behavior	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})
Cognitive	.12	73	.92	35	.02	71	.85	5
Alternate	.50	2	.60	29			.86	4
Placebo	.01	47						
Sociopathic Behavior								
Cognitive	.42	10			.55	6		
Alternate	.05	8			1.04	2		
Placebo	-.10	3						
Discrete Behaviors								
Cognitive	.00	2	,57	13	.09	12	.32	8
Alternate			1.20	11			.50	4
Placebo	.00	2	.00	4				
Impulsivity (self report)								
Cognitive	.26	11						
Alternate	.18	10						
Placebo								

Table 9 (continued)

Cognitions	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})
Cognitive	-.08	3	.36	17			.48	3
Alternate	.37	3	.23	7			.49	3
Placebo			-.77	3				

in comparisons against no treatment controls. Nearly 200 outcomes against placebo controls show only average effectiveness for cognitive restructuring, surpassed somewhat by the typical alternate treatment. A similar pattern emerges with trait anxiety measures, except that all treatments fare poorly against placebo controls. Public speaking measures show a pattern similar to that of state anxiety. The constructs measured are likely to be similar in practice, judging from the designs reviewed.

Comparisons against no treatment produced a very large mean effect for cognitive restructuring in treating depression; against placebos, therapy effected little change. Measures of other affects produced a salient mean effect for cognitive restructuring against placebos in the most substantially represented affect category for children. In this case a modest number of outcomes against no treatment revealed little effect.

Child treatment was often assessed with achievement and cognitive/performance measures. This largely reflected treatment of impulsivity, relative to school related tests, intelligence subtests, mazes, etc. (achievement measures) or task reaction times (cognitive/performance). Achievement tests consistently show little typical effect. Cognitive restructuring fared far above average for performance measures but was nonetheless surpassed by a very large mean effect for the modest number of alternate treatment outcomes against placebo controls.

For adults, achievement was generally academic improvement in college students. Cognitive tasks were again largely reaction times or information processing tasks. The former indicated a large effect for 11 outcomes of placebo versus no treatment and a typical effect for cognitive restructuring against placebos. Cognitive tasks, with a modest number of observations, reflected a major psychotherapy effect.

Another heavily tested category encompassed measures of assertiveness. Margins of greater effectiveness of cognitive restructuring are borne out across both sets of comparisons. Similarly, but in small numbers, cognitive restructuring is notably effective for psychotic behavior, and vocationally oriented measures. With less consistency across control types, cognitive restructuring also fared well in some instances for social and sociopathic behaviors, indicators of psychosocial adjustment, self concept, and somatic variables.

Across all treatments and all measures, the length of follow-up periods (lapse between treatment termination and testing) mattered little. The overall correlation of -.04, unimpacted by ES weighting, suggested negligible loss of therapeutic gain across time.

Certain psychometric qualities of outcomes proved interesting. The modality or "sphere" of measurement reflected differential mean outcome, other things being equal (see Table 10). Physiological measures showed very little gain. Behavioral measures for adults differed little from self-

Table 10

Influence of Measurement Characteristics

Variable ..Level ..Treatment	No Treatment Controls				Placebo Controls			
	Children		Adults		Children		Adults	
	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}
Measurement Sphere								
A. Behavioral								
Cognitive	.33	167	.76	166	.30	195	.65	107
Alternate	.17	36	.49	103	1.01	17	.69	64
Placebo	-.01	78	.20	49				
B. Self Report/Cognitive								
Cognitive	.58	120	.75	570	.22	114	.52	365
Alternate	.40	28	.50	217	.28	17	.37	154
Placebo	.27	34	.22	163				

Table 10 (continued)

C. Physiological	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})	$(\overline{ES})$	(N_{ES})
Cognitive			.12	26			-.01	15
Alternate			.04	10			.06	10
Placebo			.02	7				
Diagnosis Method								
A. Self								
Cognitive			.67	129			.29	47
Alternate			.67	57			-.06	3
Placebo			.21	17				
B. Psychometrics								
Cognitive	.17	132	.71	511	.30	141	.56	364
Alternate	.15	29	.48	199	.72	28	.47	180
Placebo	.03	43	.20	159				
C. Other Clinical								
Cognitive	.70	128	1.00	110	.22	112	.63	59
Alternate	.38	35	.35	66	.30	6	.29	33
Placebo	.10	42	.25	29				

report for no treatment comparisons, but showed a decisive advantage in comparisons against placebo controls. Children treated with cognitive restructuring, in contrast to placebo treated children, showed greater mean change in behavioral spheres than in congitive spheres; in comparisons with untreated children, cognitive spheres entailed the greater mean change. Alternate treatment appeared to be unusually effective in bringing about behavior change in children.

The categorization of types of disorders treated is considered further in Table 10. Diagnostic labels were generally assigned in one of three ways: (a) self-selection, according to subject requests for help or volunteering for promoted treatment projects, (b) psychometrics, where subjects exceeded some minimal pretest level, and (c) other methods of clinical assessment. Cognitive restructuring clearly performed best with subjects clinically diagnosed in some manner. Self-diagnosed subjects responded to treatment at about the average for contrasts against no treatment, but well below average against placebos.

A final concern was the level of accuracy entailed in calculating ESs. In most studies, the data reported permitted adequate ES estimation. In some, however, only rough probability cutoffs were reported, necessitating a cruder estimate with restricted range and variability. All else being equal, the distribution of ESs into the dichotomy of "exact" versus "estimated" produced point biserial correlations of .13 and .16 for comparisons against no treatment and placebo controls,

respectively. On the average, ESs more crudely estimated were smaller.

It bears comment that, aside from ES accuracy, effect size correlated substantially with the level of significance for the corresponding effect, where reported (r_{xy} = .66 (no treatment); r_{xy} = .71 (placebo)). Conversely, less than 50% of the variability in treatment effectiveness can be accounted for in the (traditional) method of considering categorical levels of significance alone.

Treatment parameters. Several variables are concerned with the format, quantity, and dispersion of therapy. Modality of treatment, in particular, delineated major sources of variability (see Table 11). Individual cognitive restructuring therapy was much more effective for adults than group therapy, and the trend was reversed for children. The 27 observations for adult treatment with mixed modality, only suggestive in their modest number of ESs, had the largest $\overline{ES}$ of no treatment contrasts. The age differences in response to modality for alternate treatment were inconsistent across comparison group sets.

The quantity and dispersion of therapy can be considered in numerous ways, such as length of sessions, span (number of weeks therapy was in progress), and total number of therapy sessions.

The most intrinsically interesting variable is probably the overall number of therapy sessions. Its correlations with outcome, correcting for ES weighting, were .01 (no

Table 11

Treatment Effectiveness by Modality

Modality ...Treatment	No Treatment Controls				Placebo Controls			
	Children		Adults		Children		Adults	
	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}
Individual								
Cognitive	.34	243	.93	306	.26	291	.71	236
Alternate	.26	39	.74	89	.72	28	.42	77
Placebo	.08	110	.24	93				
Group								
Cognitive	.99	47	.58	435	.31	18	.36	246
Alternate	.29	25	.38	221	.30	6	.46	151
Placebo	.00	2	.18	125				

Table 11 (continued)

Mixed Modality	$(\overline{ES})$....(N_{ES})	$(\overline{ES})$....(N_{ES})	$(\overline{ES})$....(N_{ES})	$(\overline{ES})$....(N_{ES})
Cognitive		1.06 27		.23 10
Alternate		.58 21		
Placebo				

treatment controls) and -.08 (placebo controls). The results for therapy span were only a little more encouraging, with correlations of .10 and -.01, respectively. Length of therapy sessions, considered alone, produced the largest correlation (r_{ES} = -.01 (no treatment controls); r_{ES} = .16 (placebos)). The CVI was inconsequential throughout.

Therapy outcome is certainly related to a multitude of factors, and one would not expect to find simple univariate relationships which account for large proportions of its variability. Nonetheless, reason would argue for a substantial relationship between the overall amount of therapy and the size of its effect. The largest relationship cited accounts for scarcely 3% of the variability. Even this is predicted from only the length of individual therapy sessions, not the overall number or dispersion of sessions. Scatterplots failed to suggest curvilinearity.

A number of other variables could be included within considerations of "experiment characteristics". A group of these which may be of interest are presented in Table 12. Generally all relationships are small and essentially unaltered by partialling out the influence of ES weighting and the CVI. Mortality had been previously considered in its contribution to the validity indicator. Viewed separately, mortality of controls and treatment groups is inversely related to outcome. Studies with high mortality, all else being equal, had somewhat lower outcome. Sample size was similarly related

Table 12

Experiment Characteristics[a]

Variable	No Treatment Controls			Placebo Controls		
	Correlation	Weight Controlled	CVI Controlled	Correlation	Weight Controlled	CVI Controlled
Treatment Mortality	-.16	-.16	-.16	-.14	-.14	-.14
Control Mortality	-.15	-.15	-.15	-.07	-.07	-.06
Sample Size (treatment)	-.04	-.04	-.04	-.15	-.16	-.16
Sample Size (control)	-.03	-.03	-.03	-.18	-.19	-.19
Coding Confidence	.00	.01	.00	-.05	-.04	-.06

[a] Values reflect simple correlations with outcome and correlations controlled for (partialling out the influence of) outcome weighting and the composite validity index (CVI).

to outcome, at least for comparisons against placebos. Larger samples yeild smaller effects, perhaps by the mechanism of minimizing "file drawer" effects. The level of detail reported in studies (classification confidence) was unrelated to outcome.

Subject characteristics. An array of subject variables was addressed in multiple regressions on ES. Analyses were produced via the Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). Entry into the regression equations was ordered by sequential contribution of variables remaining. Results are portrayed in Table 13.

The relatively low level of multiple correlation with ES across all regressions is of first concern. Roughly 5% to 9% of ES variability is predictable across equations. The potential for capitalization on chance with seven predictors argues for cautious interpretation, and an R^2 of .09 is uncomfortably low in this multivariate context. Nonetheless, small relationships are not necessarily artifacts. For descriptive purposes, interpretation of standardized Beta weights thus proceeded, but with caution.

The five regressions (with exclusive ESs) constitute partial replications in a limited sense. In this light, it is interesting to note that certain relationships emerge quite consistently throughout. Surprisingly, I.Q. was inversely related to predicted outcome in four regressions, relatively

Table 13

Subject Characteristics Regressed on Outcome[a]

Variable	No Treatment Controls			Placebo Controls	
	Cognitive	Alternate	Placebo	Cognitive	Alternate
Per cent Female	.11	.13	.11	.03	.10
IQ	-.39	-.03	-.26	-.25	.13
SES	-.03	.11	-.33	-.11	-.15
Age	-.01	.09	.05	.03	-.01
Similarity to Therapist	.38	.04	.36	.28	.07
Sex Match to Therapist	-.16	-.09	-.17	-.01	-.09
Per cent Minority	-.01	.14	-.17	-.01	.14
Multiple Correlation	.30	.28	.29	.22	.27

[a] Values represent standardized regression coefficients.

strongly in three. Similarly, increased similarity of subjects and therapists generally predicted more favorable outcome. At the same time, match of subject and therapist gender, another "dimension" of similarity, appears to have slight, if any detrimental influence. Female gender of subjects was a modest, consistent predictor of favorable outcome.

The proportion of subjects who were minorities showed little consistent impact, although this may be attributable to the relative infrequency with which this variable could be coded. Contrasts of alternate treatments showed some favor of minorities. Predicted outcome with placebos was substantially inversely related to SES and, more modestly, to minority per cent.

It is reassuring but not too surprising that regressions based on the two most similar sets of ESs (applications of cognitive restructuring) revealed a rather consistent pattern of regression weights. I.Q. exerted a dominating inverse relationship, balanced by a positive relationship between subject -therapist similarity and predicted outcome. Female gender and gender mismatched with therapists were modest predictors with no treatment controls, but of little consequence with placebos.

The weighting of ESs was inconsequential when accounted for in the predictor set. Similarly, accounting for the CVI and confidence in rating I.Q. and SES increased overall predictability by less than 1% in all cases, with no detectable

influence on the pattern of regression loadings.

The optimal pattern of subject characteristics within the present data can be illustrated by calculating the composite outcome for "ideal" circumstances where all standardized Beta weights would contribute to higher outcome. Thus, in contrast to a placebo control, a caucasian, 30 year old lower class female, in a profession but below average in intelligence, would have been expected to show an effect size of 1.46. While this is an unusual subject prototype, it is descriptive of the trends evidenced within the data collected.

Experimenter characteristics. Analogous regression procedures were used to depict the observed influences of therapist characteristics. Training orientation (psychology versus other) and treatment (self-instructional versus other) were included as dichotomous predictors. Subject-therapist similarity were again included since, by their nature, they can be viewed as either subject or experimenter characteristics.

The typical multiple correlation (see Table 14) accounted for less than 8% of outcome variability. Again, descriptive implications of regression loadings should be viewed with caution. As with subject regressions, ES weighting, CVI, and confidence ratings had no notable impact. Unlike subject regressions, there is little consistency in the predictive utility of therapist variables. The exception may be subject-therapist match on gender, which had a negative, typically salient influence on predicted therapy outcome. It is

Table 14

Therapist Characteristics Regressed on Outcome[a]

Variable	No Treatment Controls			Placebo Controls	
	Cognitive	Alternate	Placebo	Cognitive	Alternate
Sex	.17	.04	-.02	.21	.00
Experience	-.04	.12	.02	-.05	.24
Number of Therapists	-.05	-.01	-.11	-.10	.04
Therapist Training	-.10	-.11	-.03	.10	.23
Treatment Orientation	-.05	-.14	-.13	.08	.36
Similarity to Client	.07	.14	.05	-.01	.06
Sex Match to Client	-.12	-.06	-.10	-.04	-.40
Multiple Correlation	.28	.27	.19	.28	.35

[a] Values represent standardized regression coefficients.

interesting to note that subject-therapist similarity, rather potent among subject predictors, made virtually no contribution within the set of therapist predictors. Seemingly inconsistent, the implication is actually quite interpretable. Subject-therapist similarity, across all distinctions of therapists is an important predictor among subject characteristics; but across all subject types, subject-therapist similarity contributes little which is not accounted for by other, more potent therapist variables.

Cognitive restructuring, across both types of controls, shows an added consistent influence of therapist's sex. Male gender of therapist predicts a more favorable outcome. Using the full set of therapist variables one can predict the outcome for an artificial, highly favorably context: In contrasting cognitive restructuring against a placebo control, a professional female subject, treated by a psychology-trained male therapist, experienced in self-instructional therapy, would be predicted to show an effect size of .85.

Inferential Statistics. Statements of significance of findings have a limited but important role in the present meta-analysis. Mean effects, differential mean effects, and covariation are of primary interest, and stand alone in their description of the nature of observed cognitive restructuring outcomes. The contribution of inferential statements in this context aids in providing assurances that observed effects are unlikely to be reversed or discredited

by studies unlocated, unreported, or yet to be undertaken.

Various methods have been proposed by Rosenthal (1978) for producing an aggregate probability statement from the combined probabilities of unit investigations. The method appearing to receive the most support and use is the calculation of the aggregate "standard normal deviate" (Z_m). In this method (attributed to Stouffer (1949)) Z_m is calculated as:

$$Z_m = \frac{Z_{s1} + Z_{s2} + \ldots Z_{sn}}{\sqrt{n}}$$

where, assuming unit variance, $Z_{s1} \ldots Z_{sn}$ are the standard normal deviates for unit studies and n is the total number of observations. The standard normal deviate for unit investigations (Z) is simply the common Z score (of the standard normal distribution) associated with the one-tailed probability reported in the study. All probabilities must be expressed as one-tailed. Disconfirming findings thus subtract from the aggregate Z_m. The obtained Z_m is in turn located on a standard normal curve, providing an overall probability estimate for the combined unit probabilities.

An algebraic variation of the same equation provides an estimate of the number of new (unknown or file drawer) studies of null effect (Z = 0.00) which would be required to reduce Z_m to a given level:

$$X = \left(\frac{Z_{s1} + \ldots Z_{sn}}{Z_p} \right)^2 - n$$

where X is the number of null studies necessary to reduce Z_m to the new level, Z_p.

The application of both formulae follows most directly from effects for which reported probabilities were also retrievable from the unit study. This subgroup represents 82% of the ESs for no treatment controls and 78% for placebo controls. Looking at the effects of all forms of psychotherapy against no treatment controls, the mean effect of .60 is associated with a Z_m of 26.19, for a combined significance level of much less than the lowest probability calibration available, .0000001 (which corresponds to Z = 5.50). The number of null effects required to reduce this effect to a combined probability level of .05 would be roughly 304,809. Against placebo controls, with an overall mean effect of .44, Z_m = 11.08, for a combined probability again much less than .0000001. About 36,316 null effects would be required to reduce the finding to a combined probability of .05.

It is noteworthy that since use of directional versus two-tailed tests in unit studies was not always clear, the above figures assume all tests were two-tailed. Since two-tailed significances are halved to produce Z scores for each observation, the aggregate probabilities reported are most likely very conservative estimates. Furthermore, overall mean effects for the subsets of ESs with retrieved significances are under-

estimates (.51 and .36, respectively) in comparison to the full sets of all ESs.

It has been documented elsewhere (Tout & Dush, 1980) and demonstrated presently that the method of combined probabilities tends to favor rejection of the null hypothesis of "no effect" when very large numbers of observations are encompassed. The concern is paralleled in ordinary analyses, in which even small effects are "significant" with very large sample sizes. Consequently, in either primary or meta-analyses it is important to consider both the magnitude and significance of findings.

It is contrary to present priorities to attach "significance" levels to all mean effects that have been considered. It may be instructive, however, to examine the aggregate significances of the $\overline{ES}$s of the major therapeutic categories of outcomes. Table 15 illustrates that for all categories of treatment, across both sets of comparisons, probability of observance by chance are very low. Rosenthal (1978) suggests that, as a rule of thumb, a comfortable Z_m should require "5n + 10" null observations to reverse conclusions. Within Table 15, only contrasts of placebo against no treatment, by a small margin (40 observations), fails to meet this conservative criterion. Cognitive restructuring exceeds the criterion by a factor of 7:1 against placebo controls and 43:1 against no treatment.

Table 15

Inferences from Combined Probabilities

Treatment	Standard Normal Deviate	Combined Probability	Number to Reverse[a]	ESs Observed
No Treatment Controls				
Cognitive	24.24	$p < .0000001$	187,588	873
Alternate	9.09	$p < .0000001$	12,144	414
Placebo	3.98	$p < .00003$	839	174
Placebo Controls				
Cognitive	9.77	$p < .0000001$	21,281	625
Alternate	5.24	$p < .0000002$	1,801	198

[a] Number of null (0.00) effects which would need to be observed to reduce the combined probability to less than .05.

Cognitive Restructuring: A Components Analysis

A principal intent of the present investigation was to explore the parameters of observed efficacy of cognitive restructuring. Its contrast against other types of treatment certainly speaks to the issue. Additional information, however, can be obtained from examination of the determinants of outcome within the process of cognitive restructuring itself.

Differences attributable to treatment orientation have appeared in earlier analyses. These are further articulated in Table 16. Clearly, the vast majority of outcomes, particularly for children, are patterned after the self-instructional therapy of Meichenbaum (1977). While the number of ESs is imbalanced, present findings are suggestive of major areas of difference. Unfortunately, differential outcomes across type of control group complicate interpretation.

In contrasts against no treatment, expressed orientation to SIT produced substantially lower mean outcome for children and adults. Against placebo controls the orientation differences all but disappeared for children; for adults, an explicit SIT orientation yeilded larger effects, by a factor of nearly two to one. Orthodoxy of application is an important qualifier, however. For both sets of comparisons, studies in which Meichenbaum directly participated as an author produced very large mean effects throughout. Other experimenters following the procedures of SIT fared somewhat more poorly

Table 16

Cognitive Restructuring Orientations

Variable ...Level	No Treatment Controls				Placebo Controls			
	Children		Adults		Children		Adults	
	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}	$\overline{ES}$	N_{ES}
Orientation								
Meichenbaum	.38	266	.70	599	.26	285	.62	322
Other	1.15	24	.85	169	.33	24	.35	170
Orthodoxy								
Meichenbaum Authored	1.10	10	1.23	95	.96	20	1.24	41
Meichenbaum Orientation	.38	217	.62	398	.22	252	.48	255
Other	.53	59	.74	275	.21	37	.44	196

than other approaches to cognitive restructuring in contrasts to no treatment, and about equally in contrasts to placebo controls.

Aside from nominal orientations and allegiences, the composition of therapeutic procedures was delineated in a multiple classification scheme encompassing all identifiable treatment components. It should be emphasized that component categories are not independent, as any ES may have been associated with one or more components. In fact, on the average, each ES represents the influence of four separately coded treatment components. The total number of components incorporated into each treatment correlated little with ES for either children (-.07) or adults(-.01).

Each treatment component substantially represented (present for 5% of ESs in the respective set of outcomes) was incorporated in multiple regressions for outcomes. Regressions were performed separately for children versus adults and placebo versus no treatment controls. Given the larger ratio of predictors to observations, more stringent criteria (F-to-Enter $\geq$ 1.00) were adopted for inclusion of a variable in the final regression equations. Otherwise, variables were entered in order of their sequential contribution.

Regressions of treatment components on outcome for children are presented in Table 17. With no treatment controls, seven predictors entered into the regression for a multiple correlation of .32, accounting for roughly 10% of outcome

Table 17

Cognitive Restructuring Components Analysis: Children[a]

Component	No Treatment Controls			Placebo Controls		
	N_{ES}	Correlation with ES	Beta[b]	N_{ES}	Correlation with ES	Beta[b]
SIT[c]	290	.01	-.19	309	-.11	-.13
Rational Restructuring	39	.05		1	-.03	
Imagery	35	.16	-.05	0		
Covert Self-reinforcement	163	.00	.12	202	-.04	
Modeling	258	-.09	-.19	209	.01	.09
SIT Content Manipulation	60	-.09	-.22	99	-.11	
Role Playing	52	.05	.26	42	-.05	
Behavioral Rehearsal	31	-.15	-.39	9	-.18	-.16
Skill Training	24	-.05		23	-.01	
Relaxation	32	.01		0		
Other Behavioral	18	.07		272	-.07	
Other	27	-.05		8	-.18	-.14

[a] All components present in at least 5% of ESs included.

[b] Standardized regression coefficients for all variables entering the equation (minimum F-to-enter = 1.00).

[c] Self-instructional therapy is coded 2 if the principal therapeutic mechanism, 1 otherwise.

variance. Self-instruction, present in all observations, fared more favorably when secondary to other treatment components than when it was the identified principal therapy mechanism. Procedures incorporating behavioral rehearsal, modeling, and multiple variations of self-instuctional content ("content manipulation") had lower predicted outcome. Role playing and, more modestly, covert self-reinforcement lead to more favorable predicted outcome. A hypothetical combination incorporating all positively weighted procedures and omitting all negatively weighted ones would have had a predicted outcome of 2.90.

Outcome against placebo controls was more simplistically predicted. Four variables entered the equation (R^2 = .27), accounting for about 7% of ES variance. Again, self-instruction was more complimentary if used as a secondary procedure. Presence of behavioral rehearsal or the group of "other therapies" diminished predicted outcome. Modeling provided a slight edge to predicted effectiveness. The subsequent most favorable combination of procedures according to weight directionality had a predicted outcome of .60.

A comparable components analysis for adults is presented in Table 18. With no treatment controls, seven variables entered the prediction equation. The multiple correlation of .37 accounted for 14% of the ES variance. The prominence of SIT had no bearing on predicted outcome. Presence of rational restructuring, behavioral rehearsal, and "other" behavioral techniques was substantially enhancing; imagery

Table 18

Cognitive Restructuring Components Analysis: Adults[b]

Component	No Treatment Controls			Placebo Controls		
	N_{ES}	Correlation with ES	Beta[b]	N_{ES}	Correlation with ES	Beta[b]
SIT[c]	768	-.02		492	.23	.15
Rational Restructuring	452	.07	.13	343	-.20	
Imagery	174	-.16	-.18	107	.02	
Covert Self-Reinforcement	142	.07	.06	87	.14	
Modeling	359	.06		237	.14	.12
SIT Content Manipulation	151	-.04		102	.14	.09
Role Playing	77	.07		58	-.08	
Behavioral Rehearsal	152	.15	.12	82	.17	
Relaxation	93	.02	.07	61	-.12	
Other Behavioral	60	.25	.23	75	.32	.28
Other	82	.00		31	.08	
Other Cognitive	111	-.08	-.19	104	-.21	-.12

[a] All components present in at least 5% of ESs included.

[b] Standardized regression coefficients for all variables entering the equation (minimum F-to-enter = 1.00).

[c] Self-instructional therapy is coded 2 if the principal therapeutic mechanism, 1 otherwise.

and "other" cognitive approaches diminished predicted effect. The optimal combination of included/excluded procedures according to directionality of loadings for the present analysis would have been predicted to produce an ES of 1.17.

In contrasts against placebos, a much different pattern emerges. Prominence of SIT becomes a salient codirectional predictor. Rational restructuring becomes a hindrance if considered alone and irrelevant when considered in the multivariate context. Presence of "other" behavioral procedures, though meekly represented, contributed favorably and substantially. A modest contribution was made by modeling, counterbalanced by the polarized influence of "other cognitive" components. The fifth and final variable to enter the regression equation, presence of SIT content variations, made a small positive contribution, with a multiple correlation (.42) accounting for 18% of ES variance. The predicted outcome for a hypothetical combination incorporating the contributors and omitting the balance of predictors was 1.62, an optimistic figure in the context of observed outcomes against placebo controls.

The commonalities among regressions were minimal, both across and within age groups. For children, predicted outcome consistently was higher if SIT was a secondary procedure and behavioral rehearsal was absent. Predictions for adults were even less consistent. Supplemental "other" behavioral procedures aided predicted outcome throughout, and "other" cognitive

procedures detracted. Principal cognitive restructuring procedures, however, had opposing influences, in opposite directions across control group sets. This was most extreme for placebo controls, where SIT enhanced and rational restructuring, at least via its bivariate relationship to ES, detracted.

Discussion

Two dimensions of the present analyses proved to have broad substantive and methodological implications. Distinguishing outcomes by type of control group and age of subjects revealed major differences in effectiveness and its determinants. In many instances the mean effects, correlations, and regression patterns appeared to differ in important ways across these primary subgroups. It is difficult to make broad statements regarding efficacy that are equally true for child and adult therapy, or that are independent of the type of control group. Consequently, it is also difficult to make direct comparisons between present analyses and earlier meta-analyses of therapy outcome, where these dimensions were given only cursory attention.

Direct calculation of ESs from the contrast of therapy and placebo delineates variance which is in large part orthogonal or contrary to that obtained from simply contrasting therapy and placebo to a common no treatment group. The mean advantage of therapy over placebo proved roughly comparable with either method, as would be algebraically expected. The variances of ESs, however, are not interchangeable. That which determines the effectiveness of therapy over placebo treatment is not necessarily the same determinant of therapy's superiority

to no treatment at all. Indeed, the determinants differed considerably.

Inconsistent information about efficacy across the two sets of control group comparisons is not easily reconcilable. It could be argued that contrasts against placebos should be emphasized in cases of contradiction, since they are the more rigorously controlled. In a methodological sense, this is the conservative strategy. Salient findings appearing only in the contrasts against no treatment seem worthy of at least some attention, nonetheless. In some cases, though rare in present analyses, differences could mean that the methods or areas of application are not comparable across sets of comparisons. New methods may have been initially demonstrated in contrast to no treatment, and not yet replicated in a more rigorous design. In such cases, or even when comparable treatments differ across control type, areas of consideration for further research are suggested.

The problem of reconciliation of findings across rigor of control has even broader implications beyond the present analysis. Simply put, the problem could be minimized if more investigations incorporated adequate controls. Journals continue to publish, and dissertation committees continue to approve, large numbers of outcome studies without placebo group designs. Many of these in turn find their way to prominence, aided by the greater

likelihood of impressive findings under poorly controlled circumstances (considering present mean ES differences). Through general dissemination to the professional community and literary reviews which underemphasize methodological control, the field is progressively shaped by outcomes of inadequate experimental designs. Ultimately, efforts to understand and capitalize on the conditions for optimal therapy outcome can only be hindered.

It was not terribly suprising to find that child therapy differs from adult therapy. More suprising, and illustrative of the problems inherent in the multivariate context of meta-analysis, was the initial near-zero linear correlation between ES and actual subject age. Collapsing on age on the basis of this early low correlation would have ignored an array of highly significant findings. Distinctions between adults and children were nonetheless, since clinical methods and theory would suggest that differences <u>should</u> exist. In most interactive relationships subsequently examined, they did. Prioritized variables of importance in a meta-analysis bear consideration beyond initial bivariate analyses.

Perseverance does not always come to fruition. The composite validity indicator proved to be largely irrelevant to outcome. Less than a disappointment, assurance is thus provided that, aside from control group design, little collective influence of validity indicators was

apparent in a manner of systematic bias. In some cases influence was further diminished or reversed in the more methodologically sound comparisons against placebos. Not infrequently, differences which did emerge tied the presence of design weaknesses to lower mean effectiveness--a source of some consolence for those who endeavor to produce sound experimental evidence.

Certain design characteristics were particularly noteworthy in their implications. Reactivity or fakeability of measurement, a matter of serious concern in its own right, was empirically shown to be largely compensated for in the placebo group design. This bias, and perhaps other "non-specifics" which effect threat to internal validity, can apparently be accounted for, if not eliminated, by contrasting treatment effectiveness against a credible non-treatment procedure. Not only therapist attention, expectancy of improvement, and other therapeutic generic factors, but design and measurement artifacts as well can thus be controlled . Moreover, present control equivalency measures suggest that comparable credibility and expectancy factors is likely an important, but not highly sensitive concern. At best, however, not all design flaws can so easily be accounted for. The most striking example was the validity scale, for which high subject attrition, particularly if imbalanced across groups, appears to leave a more favorable picture of therapy outcome for those who remain in therapy.

Aggregates of data most comparable to Smith and Glass's (1977) meta-analysis produced areas of similar findings. Cognitive restructuring presently produced the same mean gain as that observed across all 375 Smith and Glass studies, with therapy moving the typical treated subject from the 50th percentile of the control group to the 75th. Alternate treatments as a group were less potent in comparison, including the behavioral procedures which fared very well in the former meta-analysis. In turn, placebo treatment also offered more meager gain presently over no treatment.

Comparisons against placebo controls at first blush seem to reflect mean therapy effects considerably inferior to those of earlier meta-analyses. Actually the contrary is true. Smith and Glass (1977) did not directly contrast therapy against placebo, but earlier estimates suggested that, categorically, therapy appeared to provide only a 5 percentile shift advantage relative to placebos. The presently found advantage more than triples that figure, with the typical therapy client moving from the 50th to the 68th percentile of placebo controls. With over 20,000 nonsignificant new findings necessary to reverse this conclusion, there is little reasonable doubt that therapeutic gain was achieved above and beyond placebo effects. Again, this finding is virtually identical for both placebo groups of measured equivalence and those without equivalence measures.

Perhaps placebo groups are generally well designed or our equivalency measures are generally weak.

The distribution of ESs, as noted, generally tended to be somewhat leptokurtic and positively skewed. The former is at least partly attributable to the restricted range and roughly bimodal distribution ("significant/nonsignificant") of ESs calculated from only probability statements or cutoffs reported as outcomes. Skewness reflects occasional highly favorable outcomes and relatively few negative outcomes. Highly negative outcomes were extremely rare. This may be an indication of a "file drawer" problem, a concern which present inferential procedures suggest is very unlikely to be of serious dimensions. With all parameters considered it is probably reasonable to assume that the true underlying distribution of ESs approximates a normal distribution.

The file drawer problem is addressed more informally by considering mean effects of published studies versus unpublished dissertations, a source of data more easily tapped exhaustively. On the whole, mean effects were at least 20% larger for published studies. This is almost certainly a very conservative indication; many dissertations are published. Those that remain are probably selectively more representative of the lower range of outcomes.

Most characteristics of experiments, even when considered simultaneously in related groups, predicted fairly small prcportions of outcome variance . The highly complex

multiply-determined aspects of therapy, outcome, and measurement belay hopes of strong, straightforward bivariate relationships and comfortably predictable effects. Characteristics of therapists and subjects provide cases in point. Clearly these are areas of clinical import, yet multiple regressions produced only modest multivariate prediction. Moreover, the relative contributions of variables in these groups often emerged in priorities differing from general clinical assumptions.

The relatively inconsequential influence of therapist experience is particularly noteworthy, perhaps disconcerting. Its implications are best viewed in light of the fact that very few highly experienced therapists were represented. The distribution of experience essentially assigns the all-but-dissertation doctoral candidate the rank of "highly experienced psychotherapist." It is at least a hindrance that therapists of even average experience represented in applied work are very rarely participants in psychotherapy research. Conservative tests of effectiveness, perhaps even serious problems in generalizations to the clinical setting may be the cost.

Male therapists appeared to have a slight edge, especially with female clients. Among subject characteristics, a finding of major import for cognitive restructuring was its more favorable overall impact with subjects of lower intelligence. Here and elsewhere the legendary predisposition to

therapeutic gain of the young, attractive, verbal, intelligent and social client was not supported, at times contraindicated. Matching client to treatment in the implicated subgroups of clientele may well highlight particular advantages in specific applications of cognitive restructuring.

An interesting trend is the negative relationship between outcome and year of publication. The highly favorable results cognitive restructuring enjoyed in its infancy have steadily declined. A related observation is the markedly superior results in studies conducted by the author of SIT, most of which occurred in earlier years. Many speculations are possible. The enthusiasm for a new approach might be expected to enhance therapist effectiveness. Years of evaluating and dismantling a therapy may bring increasingly sophistocated and rigorous experimental designs and analyses. In a different vein, replications in later years may have failed to recapture all of the vital therapeutic ingredients in the constellation of procedures originally developed and employed. For example, the early Meichenbaum studies are notably distinguished in their use of highly experienced therapists by present standards. These and other possible explanations bear further analysis.

Differences across schools of therapists were also difficult to interpret. The superiority of psychology-trained therapists was not readily accounted for by other major sources of outcome variability. Here too further research and analysis are needed. Explanations may lie in closer examination of

therapy process and focus, beyond the present scope of data.

Age differences, of course, were pervasive when carefully examined. The most important difference between child and adult psychotherapy was in overall effectiveness. At higher levels of aggregation the typical effectiveness of child therapy was a serious disappointment. Unburdened in separate consideration, adult cognitive restructuring widened its advantage over alternate treatments, even while the other treatments were generally showing higher gains with adults as well.

The age differences for cognitive restructuring were substantially compounded by treatment modality. No treatment controls gave indication of much greater effectiveness of group therapy for children, while individual therapy or perhaps mixed modalities produced very large therapeutic gains for adults. These optimal age-modality matches produced mean effects in the neighberhood of 1.00--an impressive shift to about the 84th percentile. Group therapy for children lost its advantage in comparisons against placebos, although only a small number of group therapy ESs were represented. Further rigorous testing of this promising avenue of child treatment amidst generally discouraging evidence is certainly critical. Individual therapy for adults increased its margin of superiority to group therapy to 2 to 1 in comparisons against placebos. This firmly demonstrated advantage calls the relative efficacy of adult group therapy with

cognitive methods into serious question. Group therapy no doubt commands appeal to the experimenter in its time efficiency. Its low average effectiveness in comparison to individual therapy argues against its overall cost-effectiveness, however. Further research applying cognitive restructuring in the format of group therapy should first address itself to its first priority, demonstrating the conditions for efficacy comparable to individual therapy.

The great bulk of research with children addressed the problem of impulsivity. It has certainly not been well demonstrated that cognitive restructuring is an effective psychotherapy in this area of application. The scenerio is distorted somewhat by the commonly large findings of effectiveness with latency or other immediate task behavior changes. Various manners of inducing the impulsive child to slow down generally succeed. Unfortunately, generalization to other behaviors is rarely pursued. Moreover, the concomitant hypothesis that the task/test performance will consequently improve finds weak support, with performance here improving little. It would appear that the impulsive child performing poorly on relevant classroom tasks, once treated with cognitive restructuring, continues to perform about as poorly, but is slower at doing it. It is not immediately clear that therapy has been of service in such instances.

Nonetheless, this area of application has met with a few highly successful studies, and a great deal of interest and

investment. The need for an effective treatment of impulsivity is undeniable. It does not seem, however, that merely more of the same will lead to desired ends, nor do minor variations in applications appear to be the solution. Unless the early, powerful effects found with this client population can be reliably replicated, the accumulated contrary evidence implies an undesirable tradeoff for the cognitive restructuring therapist and researcher, not to mention the client. Many other areas of application, scarcely studied in comparison, appear to hold promise more worthy of pursuit with cognitive restructuring. In the same vein, it may be timely to introduce new innovations in the treatment of impulsivity. If cognitive restructuring is to be pursued any further in treatment of this disorder, present findings would suggest areas for consideration. The components analysis drew particular attention to the apparently favorable supplementation of covert self-reinforcement, role playing, and perhaps modeling. Similarly, behavioral rehearsal was implicated as unfavorable in impact when considered within the full context of treatment procedures.

It is clearly indicated that self-instruction is more effective here and with other childhood disorders as a compliment to other treatment components than as the principle or exclusive technique. This has frequently not been reflected in the actual treatments tested. Notably, this point has been emphasized from the outset by Meichenbaum

(1971, particularly the import of self-reinforcement and modeling as supplements to self-instruction. The possibility is again raised that the early, more impressive findings with children may have waned, at least partly, because of insufficient attention to the finer points of the therapeutic procedures Meichenbaum proposed.

Adult treatment with cognitive restructuring enjoyed many areas of substantial efficacy. Unassertiveness and various forms of anxiety responded favorably across a broad base of studies, generally surpassing results with alternate treatments. Voluminous evidence has accumulated in these areas, to the extent that large numbers of replications with minor variations are probably no longer needed. Of course substantial innovations may still hold promise of even greater efficacy. On the other hand, the limited resources of therapy research must be balanced. Many other fertile areas for future research presently suggested should be explored.

Against this background of optimism, cognitive restructuring bluntly confronts a great deal of inconsistency of findings across control group types. Beyond simply pursuing promising trends, further research with adults directly faces the burden of documenting reliable effectiveness under soundly controlled conditions. Child therapy, of course, was unable to amass convincing evidence of efficacy against any type of control condition. Ultimately its evidence must also be evaluated under rigorous conditions of control.

Components analysis also suggests directions for further considerations with adult cognitive restructuring. A critical discrepancy warranting attention is the relative contribution of SIT and rational restructuring. Placebo controls suggest that the treated fare better when therapy places heavy emphasis on the use of self-instruction, No treatment controls suggest that presence of rational restructuring is the critical determinant. Notably, the largest single influence in both sets of comparisons was the presence of various supplemental behavioral techniques. Conversely, supplemental cognitive techniques, including imagery, diminished the likelihood of therapeutic gain.

Implications of the interfaces of cognitive and behavioral techniques are encountered pervasively, for children and adults. The interactions are complex and persistent, but not always consistent in their pattern across ages and type of control. The most favorable combinations of techniques from each group may well prove to be specific to type of problem. It certainly appears that a strategy of merely increasing the number of procedures incorporated into the treatment package would provide no assurance of improved outcome. A complimentary duality of cognitive and behavioral procedures is evident in a limited sense. Certain combinations also appear to be contraindicated. Building upon the present outcome data may allow for prescriptive combinations of procedures on an empirical basis.

The disparity between outcome trends and parameters for children versus adults may seem troublesome. It probably is not. While some parallels should exist if both represent applications of the same model of behavior change, it is in fact unlikely that uniformity would follow throughout the specific determinants of outcome. Child therapy operates within a vastly different developmental and psychosocial context than adult therapy. Moreover, the problems respectively addressed have few similarities, and nonparallel parameters of outcome should be expected. The disappointment, of course, is that for children the determinants of favorable outcome in most respects remain elusive.

Limitations. Several of the limitations of the present analysis warrant consideration. For example, little data was available to address generalization of therapy effects. Cognitive techniques were effective in many areas in producing change measured along cognitive response dimensions such as self-reports. Evidence of change on physiological measures, in contrast, was very meager. Cognitive techniques did fare well on the average in bringing about change on behaviorally oriented measures, but this is qualified by (a) the typical inclusion of complimentary behavioral treatment procedures, which may explain the apparent generalization of cognitive interventions to behavioral change, and (b) the narrowness and intrusiveness of many of the behavioral measures. An illustration of the latter concern is the frequent use of

reaction time in a specific training task as an indicator of therapeutic change for impulsive children. Improvement in such an instance provides little assurance of change in social behaviors, school performance, or other areas which are problematic for the impulsive child. Present coding systems and analyses failed to adequately address these concerns; their complexity is difficult to fully embrace within meta-analytic procedures. They clearly remain priority concerns for future therapy outcome research.

In a more general sense, there is an inherent limitation in meta-analysis. As analyses proceed from high levels of aggregation to progressively finer ones, a point is ultimately reached where meta-analytic methods become more cumbersome and restrictive than literary methods. Highly specific questions may be posed which can only be addressed by a small subset of the coded data. In such cases a careful literary approach may better suit the task. A middle ground of combined use of meta-analytics and literary methods at fine levels of aggregation may prove to be an even more powerful approach.

An additional methodological concern is the lack of independence of ESs. Since it is likely that of measures within a single study are typically more highly intercorrelated among themselves than with similar measures from other studies, it is possible for studies with very large numbers of outcomes to contribute a disproportionate

amount of redundant information to aggregations of outcome. This is of particular concern for studies which use several similar measures of the same construct, such as trait anxiety. In line with the conclusions of Smith and Glass (1977), it was deemed more acceptable to tolerate lack of absolute independence of ESs than to use a collapsed average ES per study, which would lose a great deal of information about individual outcomes (such as measurement reactivity, construct measured, etc.). Statistically partialling out the influence of ES weighting was of assistance in evaluating a limited area of such influences. It would be possible to more elaborately covary for the number and similarity of outcome measures contributed by each study within a meta-analysis. The potential for more adequately accounting for dependence among observations in a statistical fashion bears serious attention in future meta-analyses.

Finally, it should be noted that while the criteria for inclusion of studies in the present meta-analysis encompassed the bulk of clinical research in the area, several studies were excluded because of lack of control groups. Often these involved treatment of clinically relevant problems with cognitive restructuring methods, but contrasted only against an alternate type of treatment. If such research designs come into more frequent use, perhaps via increased favoring of "minimal treatment controls", it may become imperative to estimate ESs for such studies and address them directly in

meta-analyses.

Further implications. The present analyses bear a multitude of implications for future research. Nearly every analysis of study characteristics defined or alluded to an area of concern or a point of interest. In many cases the implications extended even beyond cognitive restructuring. Numerous broad issues of experimental design and measurement were addressed. While not a direct focus, suggestive evidence also emerged relevant to the alternate treatments represented.

A massive amount of information was addressed, often intricate and complex. Within its own inherent limitations meta-analysis served parsimony, reducing variation scattered across thousands of pages, objectively, to a more comprehensible set of tables and figures. Certainly much still escapes any such condensation. But an array of summary data, presumably relevant to future research, has been distilled into a form more readily consulted. Prospects of promise and applications of questionable merit have been identified.

The present investigation supports the utility of meta-analytics in psychological research. With the growing pace of new studies in psychotherapy research alone it is difficult to keep a step ahead of the evidence. As a scarce commodity, sound scientific research is too valuable to waste. Yet even the present data point to areas where the known continues to be proven, unheeded by Glass's (1976) cautions. The efficiency, thoroughness, and objectivity of meta-analysis

can aid in identifying research trends before they are pursued beyond the point of diminishing returns or abandoned without due consideration. Statistical methods for integrating research offer a promising contribution to a goal of serious dimensions: keeping science abreast of itself.

References

References

*Arnkoff, D. B. Self-statement therapy and belief therapy in the treatment of test anxiety. Unpublished doctoral dissertation, Pennsylvania State University, 1979.

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

*Barabash, C. A comparison of self-instruction training, token fading procedures, and a combined self-instruction--token fading treatment in modifying children's impulsive behaviors. Unpublished doctoral dissertation, New York University, 1978.

*Barrett, T. J. The parameters of self-instructional training. Unpublished doctoral dissertation, West Virginia University, 1973.

Barton, M. A., & Glass, G. V. Integrating studies that have quantitative independent variables. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1979.

Beck, A. Cognitive therapy and emotional disorders. New York: International Universities Press, 1976.

*Bender, N. N. Self-verbalization versus tutor verbalization in modifying impulsivity. Journal of Educational Psychology, 1976, 68, 347-354.

*Beyer, V. M. Verbal self-instruction as a modifier of impulsivity. Unpublished doctoral dissertation, Brigham Young University, 1975.

*Burns, B. J. The effect of self-directed verbal comments on arithmetic performance and activity level of urban hyperactive children. Unpublished doctoral dissertation, Boston College, 1972.

*Camp, B. W., Blom, G. E., Herbert, F., & van Doorninck, W. J. "Think Aloud" : A program for developing self-control in young aggressive boys. Journal of Abnormal Child Psychology, 1977, 5, 157-170.

*Carmichael, S. R. The use of self-instruction in an assertion training program with chronic, institutionalized patients. Newsletter for Research in Mental Health and Behavioral Sciences, 1976, 18, 20-22.

*Carmody, T. P. Rational-emotive, self-instructional, and behavioral assertion training: Facilitating maintenance. Cognitive Therapy and Research, 1978, 2, 241-253.

Cautela, J. R. Covert conditioning. In A. Jacobs and L. B. Sachs (Eds.), The psychology of private events: Perspectives on covert response systems. New York: Academic Press, 1971.

Cohen, J. Statistical power analysis for the behavioral sciences (2nd ed.). New York: Academic Press, 1977.

*Coleman, R. E. Manipulation of self-esteem as a determinant of mood of elated and depressed women. Journal of Abnormal Psychology, 1975, 84, 693-700.

Cooper, H. M. Statistically combining independent studies: A meta-analysis of sex differences in conformity research. Journal of Personality and Social Psychology, 1979, 37, 131-146.

*Denny, D., Sullivan, B., & Thury, M. Participant modeling and self-verbalization training in the reduction of spider fears. Journal of Behavior Therapy and Experimental Psychiatry, 1977, 8, 247-253.

*Drummond, D. Self-instructional training: An approach to disruptive classroom behavior. Unpublished doctoral dissertion, University of Oregon, 1974.

Dush, D. M., Tout, L., & Gabriel, R. M. Meta-analysis validity: The case for tighter controls for methodological inadequacy. Paper presented at the Annual Meeting of the Midwestern Educational Research Association, Toledo, Ohio, 1980.

D'Zurilla, T. J., & Goldfried, M. R. Problem solving and behavior modification. Journal of Abnormal Psychology, 1971, 78, 107-126.

*Eastman, B. G. Cognitive self-instruction for the control of impulsive classroom behavior: A generalization training model. Unpublished doctoral dissertation, University of Florida, 1979.

*Elder, J. P. A comparison of response acquisition and cognitive restructuring in the enhancement of social competence of college freshmen. Unpublished doctoral dissertation, West Virginia University, 1978.

Ellis, A. Reason and emotion in psychotherapy. New York: Stuart, 1962.

*Ellis, D. E. The assessment of self-instructional training in developing self-control of aggressive behavior in impulsive--aggressive boys. Unpublished doctoral dissertation, North Carolina State University at Raleigh, 1976.

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

*Finch, A., Wilkinson, M., Nelson, W., & Montgomery, L. Modification of an impulsive cognitive tempo in emotionally disturbed boys. Journal of Abnormal Child Psychology, 1975, 3, 45-52.

*Finger, R., & Galassi, J. P. Effects of modifying cognitive versus emotionality responses in the treatment of test anxiety. Journal of Consulting and Clinical Psychology, 1977, 45, 280-287.

Fiske, D. W., Hunt, H. F., Luborsky, L., Orne, M. T., Parloff, M. B., Reiser, M. F., & Tuma, A. H. Planning research on effectiveness of psychotherapy. Archives of General Psychiatry, 1970, 22, 22-32.

Fisher, R. A. Statistical methods for research workers (7th ed.). London: Oliver & Boyd, 1938.

*Fremouw, W., & Harmatz, M. A helper model for behavioral treatment of speech anxiety. Journal of Consulting and Clinical Psychology, 1975, 43, 652-660.

*Fremouw, W. J., & Zitter, R. E. A comparison of skills training and cognitive restructuring-relaxation for the treatment of speech anxiety, Behavior Therapy, 1978, 9, 248-259.

*Freundlich, A. Modeling and behavior rehearsal versus cognitive self-statement modification approaches in assertion training, Unpublished doctoral dissertation, University of Louisville, 1977.

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

*Garrison, K. R., & Brown, R. D. The effects of cognitive modification and informed teachers on communication apprehension in children. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1979.

*Genshaft, J. L., & Hirt, M. Race effects in modifying cognitive impulsivity through self-instruction and modeling. Journal of Experimental Child Psychology, 1979, 27, 185-194.

*Gioe, U. Cognitive modification and positive group experience as a treatment of depression. Unpublished doctoral dissertation, Temple University, 1975.

*Gindi, M. M. A test of a behavior rehearsal therapy and a cognitive restructuring therapy with institutionalized adolescent and pre-adolescent males. Unpublished doctoral dissertation, The Catholic University of America, 1978.

*Girodo, M., & Roehl, J. Cognitive preparation and coping self-talk: Anxiety management training during the stress of flying. Journal of Consulting and Clinical Psychology, 1978, 48, 978-989.

*Glass, C. R., Gottman, J. M., & Shurmak, S. H. Response-acquisition and cognitive self-statement modification approaches to dating-skills training. Journal of Counseling Psychology, 1976, 23, 520-526.

Glass, G. V. Primary, secondary, and meta-analysis research. Educational Researcher, 1976, 5, 3-8.

Glass, G. V. Integrating findings: The meta-analysis of research. Review of Research in Education, 1978, 5, 351-379.

Glass, G. V., & Smith, M. L. Meta-analysis of research on class size and achievement. Educational Evaluation and Policy Analysis, 1979, 1, 2-16.

*Glogower, F. D., Fremouw, W. J., & McCroskey, J. C. A components analysis of cognitive restructuring, Cognitive Therapy and Research, 1978, 2, 209-223.

Goldfried, M. Systematic desensitization as training in self-control. Journal of Consulting and Clinical Psychology, 1971, 37, 228-234.

*Goldfried, M. R., Linehan, M. M., & Smith, J. L. Reduction of test anxiety through cognitive restructuring. Journal of Consulting and Clinical Psychology, 1978, 46, 32-39.

*Gottman, J., Gonso, J., & Shuler, P. Teaching social skills to isolated children. Journal of Abnormal Child Psychology,

1976, 4, 179-196.

*Graziano, A. M., & Mooney, K. C. Family self-control instruction for children's nighttime fear reduction. Journal of Consulting and Clinical Psychology, 1980, 48, 206-213.

*Gresen, R. The effects of instruction and reinforcement versus a multifaceted self-control procedure on the modification and generalization of behavior in schizophrenics. Unpublished doctoral dissertation, Bowling Green State University, 1974.

*Guralnick, M. J. Solving complex perceptual discrimination problems: Techniques for the development of problem solving strategies. American Journal of Mental Deficiency, 1976, 81, 18-25.

*Hahnloser, R. M. A comparison of cognitive restructuring and progressive relaxation in test anxiety reduction. Unpublished doctoral dissertation, University of Oregon, 1974.

*Harris, C., & McReynolds, W. Semantic cues and response contingencies in self-instructional control. Journal of Behavior Therapy and Experimental Psychiatry, 1977, 8, 15-17.

*Hauserman, N, Miller, J. S., & Bond, F. T. A behavioral approach to changing self-concept in elementary school children. Psychological Record, 1976, 26, 111-116.

*Hazel, L. F. Modifying impulsive responding in math: Comparative effects of self-instruction, reinforcement, and traditional tutoring approaches. Unpublished doctoral dissertation, State University of New York at Stony Brook, 1975.

Hearold, S. L. Meta-analysis of the effects of television on social behavior. Paper presented at the meeting of the American Educational Research Association, San Francisco, 1979.

*Heath, B. L. Application of verbal self-instructional procedures to classroom behavior management. Unpublished doctoral dissertation, University of Minnesota, 1978.

*Higa, W. R. Self-instructional versus direct training in modifying children's impulsive behavior. Unpublished doctoral dissertation, University of Hawaii, 1973.

*Holroyd, K. A. Cognition and desensitization in the group treatment of test anxiety. Journal of Consulting and Clinical Psychology, 1976, 44, 991-1001.

*Holroyd, K., & Andrasik, F. Coping and the self-control of chronic tension. Journal of Consulting and Clinical Psychology, 1978, 46, 1036-1045.

*Holroyd, K. A., Andrasik, F., & Westbrook, T. Cognitive control of tension headache. Cognitive Therapy and Research, 1977, 1, 121-133.

Hunter, J. E. Cumulating results across studies: A critique of factor analysis, canonical correlation, MANOVA, and statistical significance testing. Paper presented at the Annual Meeting of the American Psychological Association, New York, 1979.

*Hurd, S. E. Behavioral instruction vs. cognitive restructuring in an individualized assertion training program. Unpublished

doctoral dissertation, University of Notre Dame, 1977.

*Hussian, R. A., & Lawrence, P. S. The reduction of test, state, and trait anxiety by test-specific and generalized stress inoculation training. Cognitive Therapy and Research, 1978, 2, 25-37.

*Jacobs, P. C. Modifying impulsive behavior of second grade boys through cognitive self-instruction: Intensity of training related to generality and long range effects. Unpublished doctoral dissertation, Wayne State University, 1974.

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

*Kaplan, R. M., McCordick, S. M., & Twitchell, M. Is it the cognitive or the behavioral component which makes cognitive-behavior modification effective in test anxiety? Journal of Counseling Psychology, 1979, 26, 371-377.

Kazdin, A. E., & Wilcoxon, L. A. Systematic desentization and nonspecific treatment effects: A methodological evaluation. Psychological Bulletin, 1976, 83, 729-773.

*Kendall, P. C., & Finch, A. J. A cognitive-behavioral treatment for impulsivity: A group comparison study. Journal of Consulting and Clinical Psychology, 1978, 46, 110-118.

*Kendall, P. C., & Wilcox, L. E. Cognitive-behavioral treatment for impulsivity: Concrete versus conceptual training in non-self-controlled problem children. Journal of

Consulting and Clinical Psychology, 1980, 48, 80-91.

*Kendall, P. C., Williams, L., Pechacek, T. F., Graham, L. E., Shisslak, C., & Herzoff, N. Cognitive-behavioral and patient education interventions in cardiac catheterization procedures: The Palo Alto Medical Psychology Project. Journal of Consulting and Clinical Psychology, 1979, 47, 49-58.

*Kirkpatrick, P. W. The efficacy of cognitive behavior modification in the treatment of depression. Unpublished doctoral dissertation, University of Texas at Austin, 1977.

Kulik, J. A., Kulik, C. C., & Cohen, P. A. A meta-analysis of outcome studies of Keller's Personalized System of Instruction. American Psychologist, 1979, 34, 307-318.

*Labouvie-Vief, G., & Gonda, J. Cognitive strategy training and intellectual performance in the elderly. Journal of Gerontology, 1976, 31, 327-332.

*Lavigne, J. V. The relative efficacy of cognitive-behavior rehearsal and systematic desensitization in the treatment of test anxiety. Unpublished doctoral dissertation, University of Texas at Austin, 1974.

Ledwidge, B. Cognitive behavior modification: A step in the wrong direction? Psychological Bulletin, 1978, 85, 353-375.

*Lefebvre, R. C. Skills acquisition and cognitive restructuring operations in training assertive behaviors. Unpublished masters thesis, North Texas State University, 1979.

*Lent, R. W. A comparison between cue-controlled desensitization, cognitive restructuring, and a credible placebo in

alleviating public speaking anxiety. Unpublished doctoral dissertation, Ohio State University, 1979.

Light, R. J., & Smith, P. V. Accumulating evidence: Procedures for resolving contradictions among different studies. Harvard Educational Review, 1971, 41, 429-471.

*Linehan, M. M., Goldfried, M. R., & Goldfried, A. P. Assertion therapy: Skills training or cognitive restructuring. Behavior Therapy, 1979, 10, 372-388.

*Lowe, R. M. The effect of modeled cognitive restructuring on the frequency of breast self-examination in women. Unpublished doctoral dissertation, Georgia State University, 1977.

*Machum, J. B. A comparison of the effectiveness of assertion training alone and in combination with cognitive self-instruction training with chronic hospitalized patients. Unpublished doctoral dissertation, Rutgers University, 1976.

Mahoney, M. J. Cognition and behavior modification. Cambridge, Mass.: Ballinger, 1974.

Mahoney, M. J., & Arnkoff, D. B. Cognition and self-control therapies. In S. L. Garfield and A. E. Bergin (Eds.), Handbook of psychotherapy and behavior change. New York: Wiley, 1978.

*Malkiewich, L. E. Comparative effectiveness of cognitive restructuring and systematic desensitization with socially anxious subjects of diverse conceptual level. Unpublished doctoral dissertation, University of Notre Dame, 1979.

*Margolis, R. B., & Shemberg, K. M. Cognitive self-instruction

in process and reactive schizophrenics: A failure to replicate. Behavior Therapy, 1976, 7, 668-671.

*May, R. The treatment of test anxiety by cognitive modification: An examination of treatment components. Unpublished doctoral dissertation, University of Kansas, 1975.

*McCordick, S. M., Kaplan, R. M., Finn, M. E., & Smith, S. H. Cognitive behavior modification and modeling for test anxiety. Journal of Consulting and Clinical Psychology, 1979, 47, 419-420.

*McDonough, T. J. A comparison of self-instruction and play group therapy in lessening impulsivity and subsequent hyper-aggressivity in fourth, fifth, and sixth grade boys. Unpublished doctoral dissertation, Boston University School of Education, 1979.

McGaw, B., & White, K. R. Meta-analysis of empirical research. Workshop presented at the Annual Meeting of the American Educational Research Association, Boston, 1980.

*Meichenbaum, D. The nature and modification of impulsive children. Paper presented at the meeting of the Society for Research in Child Development, Minneapolis, 1971.

*Meichenbaum, D. Cognitive modification of test anxious college students. Journal of Consulting and Clinical Psychology, 1972, 39, 370-380.

Meichenbaum, D. Cognitive-behavior modification. New York: Plenum Press, 1977.

*Meichenbaum, D, & Cameron, R. Stress inoculation: A skills

training approach to anxiety management. Unpublished manuscript, University of Waterloo, 1972.

*Meichenbaum, D., & Cameron, R. An examination of cognitive and contingency variables in anxiety relief procedures. Unpublished manuscript, University of Waterloo, 1973.

*Meichenbaum, D., & Cameron, R. Training schizophrenics to talk to themselves: A means of developing attentional controls. Behavior Therapy, 1973, 4, 515-534.

*Meichenbaum, D., Gilmore, J. B., & Fedoravicius, A. Group insight versus group desensitization in treating speech anxiety. Journal of Consulting and Clinical Psychology, 1971, 36, 410-421.

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

*Mendoca, J. D., & Siess, T. F. Counseling for indecisiveness: Problem solving and anxiety-management training. Journal of Counseling Psychology, 1976, 23, 339-346.

Miller, T. I. Comparative effects of drug therapy, psychotherapy, and drug plus psychotherapy: A meta-analysis. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1979.

*Moore, S., & Cole, S. Cognitive self-mediation training with hyperkinetic children. Bulletin of the Psychonomic Society, 1978, 12, 18-20.

*Morris, N. E. A group self-instruction method for the treat-

ment of depressed outpatients. Unpublished doctoral dissertation, University of Toronto, 1975.

*Moss, V. L. The modification of impulsive behavior with learning disabled boys: A verbal self-instructional approach. Unpublished doctoral dissertation, University of Rhode Island, 1979.

*Mygdal, W. K. The acquisition of stress management skills by student teachers: An outcome study of stress inoculation and anxiety management. Unpublished doctoral dissertation, Baylor University, 1978.

*Nelson, W. J. The role of self-instruction and self-reinforcement in modification of impulsivity, and possible effects on mathematics and reading achievement. Unpublished masters thesis, University of Louisville, 1975.

*Neunuebel, P. M. The differential impact of systematic desensitization and cognitive behavior modification upon internal and external test anxious subjects. Unpublished doctoral dissertation, University of Missouri-Columbia, 1977.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. SPSS: Statistical package for the Social sciences (2nd ed.). New York: Mc Graw-Hill, 1975.

*Norman, W. The efficacy of cognitive restructuring versus skills training in generating behavior change. Unpublished doctoral dissertation, Pennsylvania State University, 1974.

*Novaco, R. W. Threatment of chronic anger through cognitive and relaxation controls. Journal of Consulting and Clinical

Psychology, 1976, 44, 681.

*Novince, L. C. The contribution of cognitive restructuring to the effectiveness of behavior rehearsal in modifying social inhibition in females. Unpublished doctoral dissertation, University of Cincinnati, 1977.

*Nye-Warren, S. L. Self-instructional stress management training: A comparison of the effects of induced affect and covert modeling in a cognitive restructuring treatment program for test anxiety. Unpublished doctoral dissertation, University of Washington, 1975.

*Nystedt, G. A. Self-monitoring and self-instruction: Therapeutic procedures for the treatment of depression in college students. Unpublished doctoral dissertation, West Virginia University, 1976.

*O'Brien, G. T. A component analysis of the semantic content of self-statements in the cognitive-behavioral treatment of frequent tension in college students. Unpublished doctoral dissertation, University of Iowa, 1978.

O'Leary, K. D., & Borkovec, T. D. Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research. American Psychologist, 1978, 33, 821-830.

*Palkes, H., Stewart, M., & Freedman, J. Improvement in maze performance of hyperactive boys as a function of verbal training procedures. Journal of Special Education, 1972, 5, 337-342.

*Palkes, H., Stewart, M., & Kakana, B. Porteus maze performance

after training in self-directed verbal commands. Child Development, 1968, 39, 817-826.

Paul, G. L. Insight versus desensitization in psychotherapy. Stanford, Calif.: Stanford University Press, 1966.

Paul, G. L. Outcome of systematic desensitization II: Controlled investigations of individual treatments, techniques, variations, and current status. In C. Franks (Ed.), Behavior therapy: Appraisal and status. New York: McGraw-Hill, 1969.

Pearson, E. S. The probability integral transformation for testing goodness of fit and combining tests of significance. Biometrika, 1938, 30, 134-148.

*Robinson, J. C. Efficacy of operant vs. cognitive behavior modification in the development of self-control for aggresive adolescent boys. Unpublished doctoral dissertation, University of Virginia, 1978.

*Rogers, T. The cognitive and psychophysiological effects of therapeutic self-statements: An investigation of the self-persuasion hypothesis. Unpublished doctoral dissertation, Pennsylvania State University, 1979.

Rosenthal, D., & Frank, J. D. Psychotherapy and the placebo effects. In C. F. Reed, I. E. Alexander, & S. S. Tomkins (Eds.), Psychotherapy: A source book. Cambridge, Mass.: Harvard University Press, 1958.

Rosenthal, R. Experimenter effects in behavioral research (2nd ed.). New York: Irvington, 1976.

Rosenthal, R. Combining results of independent studies.

Psychological Bulletin, 1978, 85, 185-193.

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

*Rychtarik, R. G. The reactivity of self-monitoring and self-instructional training in the modification of speech disfluencies in speech anxious college students. Unpublished doctoral dissertation, University of Montana, 1979.

*Sage, C. K. Using self-instructional techniques to teach children positive assertion. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill, 1979.

*Sarason, I. Test anxiety and cognitive modeling. Journal of Personality and Social Psychology, 1973, 28, 58-61.

*Schlichter, K. J., & Horan, J. J. Effects of stress inoculation training on the anger management skills of institutionalized juvenile delinquints. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1979.

*Shaw, B. F. Comparison of cognitive therapy and behavior therapy in the treatment of depression. Journal of Consulting and Clinical Psychology, 1977, 45, 543-551.

*Sinclair, C. Cognitive versus social-behavioral training in the generalization of a self-instructional program for impulsive children. Unpublished masters thesis, University of Calgary, 1976.

Smith, M. L. Sex bias in counseling and psychotherapy. Psychological Bulletin, 1980, 87, 392-407.

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

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

*Snyder, J. J., & White, M. J. The use of cognitive self-instruction in the treatment of behaviorally disturbed adolescents. Behavior Therapy, 1979, 10, 227-235.

Spivak, G., & Shure, M. Social adjustment of young children: A cognitive approach to solving real-life problems. San Francisco: Jossey-Bass, 1974.

Suinn, R. M., & Richardson, F. Anxiety management training: A nonspecific behavior therapy program for anxiety control. Behavior Therapy, 1971, 2, 498-510.

*Taylor, F. G., & Marshal, W. L. Experimental analysis of a cognitive-behavioral therapy for depression. Cognitive Therapy and Research, 1977, 1, 59-72.

*Terrill, M. J. The effects of assertion training and cognitive restructuring on the assertive behavior and situational discomfort of college women. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, 1978.

*Thorpe, G. L. Desensitization, behavioral rehearsal, self-instructional training, and placebo effects on assertive-refusal behavior. European Journal of Behavior Analysis and Modification, 1975, 1, 30-44.

*Tolbert, H., Jr. Self-instructional training of test anxious

students differing on locus of control. Unpublished doctoral dissertation, American University, 1979.

Tout, L. R., & Dush, D. M. A comparison of the effects of biasing factors on three statistical methods for integrating studies: Meta-analysis, voting, and combining probability levels. Paper presented at the Annual Meeting of the Midwestern Educational Research Association, Toledo, Ohio, 1980.

*Wein, K. S., & Nelson, R. O. The effect of self-instructional training on arithmetic problem solving skills. Unpublished manuscript, 1977.

*Wein, K. S., Nelson, R. O., & Odom, J. V. The relative contributions of reattribution and verbal extinction to the effectiveness of cognitive restructuring. Behavior Therapy, 1975, 6, 459-474.

*Weinreich, R. J. Inducing reflective thinking in impulsive emotionally disturbed children. Unpublished masters thesis, Virginia Commonwealth University, 1975.

*Weissberg, M. A comparison of direct and vicarious treatments of speech anxiety: Desensitization with coping imagery, and cognitive modification. Behavior Therapy, 1977, 8, 606-620.

White, K. R. The relationship between socioeconomic status and academic achievement. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1979.

*Wilcoxon, L. A. Therapeutic effects of self-instructional training for nonassertive women. Unpublished doctoral

dissertation, Pennsylvania State University, 1976.

*Williams, D. Y., & Akamatsu, T. J. Cognitive self-guidance training with juvenile delinquints: Applicability and generalization. Cognitive Therapy and Research, 1978, 2, 285-288.

*Wolf, J., & Fodor, I. Modifying assertive behavior in women: A comparison of three approaches. Behavior Therapy, 1977, 8, 567-574.

*Workman, J. V. Effects of self-instructional training upon test taking behavior of children from regular classes. Unpublished masters thesis, California State University, 1977.

*Yorde, B. S. A comparison of the effects of using cognitive restructuring, relaxation techniques, and frontalis electromyograph biofeedback training to induce low arousal response to stress. Unpublished doctoral dissertation, Ohio University, 1977.

*Studies represented in the meta-analysis.

Appendix A

Data Collection Sheets

Cognitive Behavior Therapy Outcomes

Value . Variable (code once for each study)

___ Study ID (First author:_____)
___ Record Number
___ Document Date
___ Source: (1) Journal (2) Book (3) Dissertation
(4) Other
___ Training of Author (1st): (1) Psychology
(2) Education (3) Other
___ Study called "analogue"? (1) Yes (2) No
___ Blinding: (1) E did therapy (2) E knew group
composition but didn't do therapy
(3) single blind (4) other:_____

-Subject Characteristics-

___ Primary Problem:
(1) Neurotic (2) Simple phobia
(3) Complex Phobia (4) Affect/depression
(5) Normal (None) (6) Character Disorder
(7) Delinquent/Criminal (8) Substance Dependency
(9) Somatic (10) Handicapped
(11) Mixed (12) Impulsivity
(13) Obesity (14) Insomnia
(15) Other:_____
(16) +...see codebook

___ Diagnosis Method: (1) Self (2) Extreme Score
(3) Psychometrics (4) Clinical
(5) Other:_____
___ Type of Phobia: (1) Speech (2) Tests (3) Reptile
(4) Rodent (5) Insect (6) Agoraphobia
(7) Other:_____
___ Length of hospitalization (average months)
___ IQ: (1) MR (2) Low (3) Average (4) High (5) ?
___ ---Confidence: (1) Low (2) Estimate (3) Stated
___ SES: (1) Low...(3) High (4) ?
___ ---Confidence: (1) Low (2) Estimate (3) Stated

___ % Female
___ % Minority (if stated)
___ Mean Age (nearest year)
___ Students? (1) yes (2) no

-Design Characteristics-

___ Subject Assignment: (1) Random (2) Matching (3) Pretest Equating (4) Convenience (5) Multi-stratified (6) Other:_____

___ Subject Solicitation: (1) Autonomously Presented (2) Response to Ad (3) Referred (4) E Solicited (5) Committed (6) Other:_____

___ Therapist Assignment: (1) Random (2) Matching (3) Non-random (4) Single Therapist (5) Fully Crossed (6) Other:_____

___ Counterdemand? (1) yes
___ Internal Validity (1) Low...(3) High
___ Overall Mortality: Treatment(s) (%)
___ Overall Mortality: Placebo(s) + Control(s) (%)
___ Are all groups of same modality? (1) yes (2) no
___ Number of Comparison Groups
___ Number of CBT Groups
___ Number of Placebos
___ Number of Non-CBT Treatments
___ Max # of Outcomes within Comparisons

-Treatment Characteristics-

___ CBT Orientation: (1) Meichenbaum (2) Beck (3) Ellis (4) Other:_____

___ Orthodoxy: (1) Low...(3) High

___ === Confidence of Treatment Classification: === (1) Low...(3) High

___ Equivalence Assurance Method: (1) None (2) ? (3) By Design (4) Collateral Measurement (5) Concurrent Measurement (6) Post Hoc (7) Other:_____

___ Number of principal comparisons

___ Second Treatment Type: (1) CBT (2) RET
(3) Other Cognitive (4) Education
(5) Counseling (6) Psychodynamic
(7) Client Centered (8) Gestalt
(9) Sys. Des. (10) Other Behavioral:_____
(11) Relaxation (12) Other:_____

___ Number of therapists

Comparisons Present:
___ -Second CBT
___ -Second Treatment
___ -Nonspecific Treatment
___ -Placebo
___ -Intact Group
___ -Wait List
___ -No Treatment
___ -Other:_____
_____

___ Location of Treatment: (1) School (2) Hospital
(3) CMHC (4) Other Clinic (5) Private
(6) College Clinic (7) Residential Facility
(8) College Lab (9) Prison (10) Unknown
(11) Other:_____

COMMENTS:_____
_____
_____

-Outcomes-

Value — Variable (code once for each outcome)

___ Study ID

___ Record Number

___ Comparison Number

___ # of Outcomes for this Comparison

___ Number of this Outcome

(*Name of Measure:_____)

___ "Treatment" for this comparison:
(1) Principal CBT (2) Other CBT
(3) Second Treatment (4) Placebo
(5) Nonspecific Treatment
(6) Other:_____

___ ---Type of second treatment: (see code)

___ "Control" for this comparison:
(1) Other CBT (2) Second Treatment
(3) Nonspecific Tmt. (4) Placebo
(5) Intact Group (6) Wait List
Other:_____

___ Treatment Modality: (1) Individual (2) Group
(3) Family (4) Mixed (5) ? (6) Other:_____

___ Control Modality: (see code--same as above)

_____ Primacy: (see codebook)

___ *Outcome Type: (1) Fear/anxiety
(2) Self-concept (3) Global Adjustment
(4) Adjustment Indicator (5) Personality Traits
(6) Somatic (7) Addiction
(8) Sociopathic Behavior (9) Social Behavior
(10) Discrete Behavior (11) Trait Anxiety
(12) Psychotic Behavior (13) Achievement: School
(14) Cognitive Task (15) Achievement: Other
(16) Vocational (17) Physiological
(18) Public Speaking (19) Assertiveness
(20) Depression (21) Other:_____
*(last digit: 1=behavioral, 2=self rep, 3=physio, 4?)

___ Measurement Reactivity: (1) Low...(5) High

___ Number of days after treatment measure taken

___ Measurement Reliability (if stated)

___ Reliability Type: (1) inter-rater (2) split-half
(3) test-retest (4) referenced (5) other:

___ Treatment Group Sample Size

___ Control Group Sample Size

___ Allegiance of Therapist to Treatment: (1) + (2) -
(3) 0 (4) Unknown

___ Allegiance of Therapist to Control: (same code)

___ Experience of Treatment Therapists in Years

___ ---Confidence: (1) Low (2) Estimated (3) Stated

___ Experience of Control Therapists in Years

___ ---Confidence: (1) Low (2) Estimated (3) Stated

___ Treatment Number of Sessions

___ Treatment Duration in Hours

___ Treatment Span in weeks

___ Control Number of Sessions

___ Control Duration in Hours

___ Control Span in Weeks

___ Equivalence: (1) Low...(3) High

___ Treatment Equivalence Rating

___ Control Equivalence Rating

___ Effect Size Calculation: (1) Standard (2) MS Within
(3) MS Total - Treatment (4) Probability
(5) Chi Square (6) T (7) F (8) Mean & p
(9) Nonparametrics (10) Correlations (11) Raw
(12) Estimates (13) % (14) Other:_____

___ Source of Means: (1) Post Test (2) Adjusted
(3) Residual Gains (4) Pre-Post (5) Other:

___ Significance of Effect: (1) -.001 (2) -.01 (3) -.05
(4) -.10 (5) N/S (6) .10 (7) .05 (8) .01 (9) .001

___ Treatment Therapist Sex: (1) Female (2) Male (3) ?

___ Control Therapist Sex: (same code)

___ Treatment Therapist/Client Similarity:
(1) Low...(4) High

___ Control Therapist/Client Similarity
(1) Low...(4) High

___ Treatment Mortality

___ Control Mortality

___ Treatment % Improved

___ Control % Improved

___ % Female - Treatment

___ % Female - Control

___ Treatment Pre Mean
___ Treatment Pre Standard Deviation
___ Treatment Post Mean
___ Treatment Post Standard Deviation

___ Control Pre Mean
___ Control Pre Standard Deviation
___ Control Post Mean
___ Control Post Standard Deviation

___ T Statistic
___ F Statistic
___ MS Within/Common/Residual
___ Effect Size (Notes:_____
_____
_____)

___ Check if another study from same report exists
Other Effects Addressed Within Study (sig = ± .10)
___ Sex: (2) Yes, N/S (3) +sig (4) -sig
___ Race: (same code)
___ Crd. (same code) - credibility
___ Age: (same code)
___ Exp. (same code) - expectancy
___ Therapist: (same code)
___ Other: (same code)-_____
___ Number of ANCOVA Covariates
___ ---Number Sig (.10)
___ Number of Factorial Effects (Excluding Treatment)
___ ---Number Sig (.10)
___ ---Probability Adjusted? (2) Yes

___ Last Effect for this Comparison? (1) Yes
___ ---If Yes: Average Effect for this Comparison
___ Last Comparison for this Study? (1) Yes
___ ---If Yes: Average Effect en toto

COMMENTS:_____
_____
_____

Appendix B

Elaboration of Scoring Criteria

Variable	Scoring Criteria
I.Q.	(...if not stated)
very low	-only if mental retardation is stated
low	-state facilities; special education classes
medium	-default (general population)
high	-college students, professions, etc.
S.E.S.	
low	-state facility, etc.
medium	-default
high	-professionals; private college, clinic, etc.
Validity	(using the highest group mortality observed:)
low	-high mortality (over 25%)
medium	-medium mortality (over 10%)*
high	-low mortality (under 10%)*
	*score lowered by 1 if nonrandom assignment
Therapist Allegiance	(proved infrequently codable)
positive	-evidence of affiliation with cognitive techniques
negative	-evidenced affiliation with alternate techniques
Therapist Experience	...estimated years (if not stated)
0	-no evidence of substantial therapy training
1	-pre-Masters graduate student (related field)
2	-Masters level
3	-post-Masters, pre-doctoral.
4	-recent doctorate; post-doctoral intern
5	-evidenced post-doctoral experience

Therapist-Client Similarity	(according to client characteristics)
low	-pre-adolescents; institutionalized; mentally retarded; chronic patients, etc.
medium	-adolescents; minorities (unless therapist is clearly of the same minority)
high	-adults in general
very high	-professionals; college students, etc.
Measurement Reactivity	(fakability)
low	-physiological measures, grade point average, other unobtrusive data, etc.
medium	-blindly scored tests (projective, behavioral)
high	-standard questionnaires, personality tests (objective), I.Q. tests, etc.
very high	-therapists ratings, non-blind projectives, teacher ratings, subject reports to therapist, etc.