1994

A Meta-Analysis of Bibliotherapy Studies

Richard Welton Marrs
Loyola University Chicago

LOYOLA UNIVERSITY OF CHICAGO

A META-ANALYSIS OF BIBLIOTHERAPY STUDIES

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF COUNSELING AND

EDUCATIONAL PSYCHOLOGY

BY

RICHARD WELTON MARRS

CHICAGO, ILLINOIS

MAY, 1994

ACKNOWLEDGEMENTS

The author wishes to extend his sincere gratitude to his advisor and the Director of his dissertation committee, Dr. Steven D. Brown. Dr. Brown has been a paragon of what it means to be a researcher. His philosophy that a dissertation should be an independent research project is greatly appreciated. The author also wishes to thank the other members of the committee, Dr. Gloria Lewis and Dr. Joseph Durlak, for their guidance and support in the completion of this dissertion and throughout my doctoral program.

Special thanks go to Laura Marrs and (now Dr.) John Gibson for their many hours of work in coding these studies to provide interrater agreement and reliability data. Both were indispensable in providing encouragement and ideas.

Sincere gratitude is also accorded to numerous administrators and faculty at Concordia University for their expressions of interest and support during this dissertation process.

Finally, loving thanks are owed to Laura, Brittany, and Kaellyn for all their sacrifices and support during this dissertation process.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Early Opinions about Bibliotherapy

Reading books for personal development has been around as long as there have been books. The inscription above the doors of the ancient library of Thebes reportedly read the "Healing place of the soul;" at Alexandria's library it said "The medicine of the mind" (Schulties, 1972). In the 20th century the process of reading for healing has come to be titled "bibliotherapy." Its acceptance by professionals as a way of helping people to help themselves has become widespread. One need only to go to a local bookstore to become overwhelmed with the number of self-help books available for many different types of problems. Early in the genesis of bibliotherapy many prominent professionals appear to have been in agreement about its positive nature (Burt, 1973). However, this acceptance seems to have been rather uncritical and based upon clinical experience rather than empirical evidence. Karl and William Menninger are frequently cited as strong early supporters of bibliotherapy (e.g., Schultheis, 1972; Brown, 1975; Burt, 1973) .

It has been estimated that over 2,000 self-help books are published each year (APA, 1989). Some of the major names in psychology and psychiatry who have written self-help books and

booklets include Carl Rogers (Rogers & Stevens, 1967), Albert Ellis (Ellis & Knauf, 1977), Aaron Beck (1989), Masters and Johnson (1970), Erich Fromm (1956), and Phillip Zimbardo (1977).

Various explanations have been given for why bibliotherapy might be effective. In a rather comprehensive review of this literature prior to 1970, Burt asserted that the majority of those who had used and/or written bibliotherapy were in agreement that it could help in "(a) achieving (emotional and intellectual) insight..., (b) verbalizing problems..., (c) externalizing problems... (d) identifying with a character or experience so that a subsequent abreaction may be achieved, and (e) thinking more constructively...." (Burt, 1973, p. 4).

The use of bibliotherapy has long been an interdisciplinary one. Librarians, social workers, psychologists, physicians, and nurses have all written about bibliotherapy in their professional literatures (Burns, 1992). Librarians have been especially prominent in the development of bibliotherapy. The use of library services for the treatment of mental patients was part of the era of American psychiatry known as "moral treatment" after the 1830's. After World War I there was a widespread adoption of bibliotherapy in Veterans Administration hospitals and numerous articles were written by VA librarians (e.g., Pomeroy, 1927; Peterson-Delaney, 1938; Kinney, 1946). In 1964 the National Institute of Mental Health funded an interdisciplinary workshop sponsored by the American Library Association entitled "Bibliotherapy: What it is and can do for

mental health" (Beatty, 1964). Over half of the non-empirical citations found in the initial literature search for this meta-analysis were from the professional literature of librarians.

In his 1969 Presidential Address to the American Psychological Association, George Miller exhorted professional psychologists to "give psychology away," i.e., to teach psychological techniques of practical importance to the general public (APA, 1978). Bibliotherapy would appear to be one way of doing this. However, the proliferation of this "giving away" (or selling in a different mode) may have outstripped our current empirical knowledge of the usefulness of bibliotherapy.

The debate about the usefulness of self-help treatment books has been an important one to the American Psychological Association for at least 15 years. In 1978 the APA appointed a Task Force to report on self-help therapies. The Task Force pointed out psychologists were in a unique position to contribute to the self-help movement because they were the only professional group that combined clinical and research experiences (Rosen, 1987). They also pointed out that a visit to any local bookstore would unfortunately demonstrate numerous member violations of APA Ethical Standards Principle 4 about fairly portrayed public statements without exaggeration or sensationalism and Principle 2 about provision of services that meet recognized standards (APA, 1978). Often psychologists allow jacketcovers and promotional blurbs to be contractually controlled by publishers who are not

university students, who alcohol educator Gerardo Gonzalez views as a "population at risk" for substance abuse, and the student culture, which he labels a "microcosm of society"? Is a "healthy student community" concept appropriate for college drinking groups (Burns, 1989)? The fact that these questions continued to go unanswered while alcohol and drug prevention programs multiplied was what interested me in conducting this study.

What I propose to discuss in the chapters ahead are: what models of prevention programs are found on university campuses; what generates student violations of prevention policy and how frequently do they occur; which legal, social and academic problems are associated with the alcohol and drug use which appears in the college student population; and what peer social group involvement generates and inhibits these behavioral problems.

Students are classified according to the strength of attachment to the "student culture" or to a pro-social culture critical of alcohol and drug use. A pro-social culture emerges when prevention awareness levels increase. Those groups, one which maintains a "party subculture" and one which develops the drug and alcohol-free culture, will be contrasted. The general point is to give meaning to the framework of change which is found with the new focus on "drug-free" campuses. One of the strategies to reach the turning point for "drug-free" schools is to promote the pro-

approval, but that his complaints to the publisher resulted in a later edition with nonextravagant claims. Holtje (1988) warned that if psychologists become too conservative in their production of self-help materials, the resulting marketplace vacuum will be filled by authors with less knowledge and experience.

The Purpose

The purpose of this dissertation is to meta-analytically study previous research about bibliotherapy, making inferences about its efficacy across various populations (e.g., various client problems) and determining its efficacy as an adjunct to traditional psychotherapy. I believe that this type of information has important ramifications for the field of counseling psychology and for our increasingly self-help oriented society in general. It may also be pertinent to the country's current struggle with health care costs and delivery systems. Bibliotherapy, if it is effective with some problems or clientele, could be used as a low-cost alternative to psychotherapy and health education.

At present there is much debate in the field (e.g., Craighead, McNamara, and Horan, 1984; Glasgow and Rosen, 1978) about the effectiveness of self-help oriented interventions: Which problems are most amenable to bibliotherapy? Which are least amenable? How much can it be expected to help? What type of person would be most helped by bibliotherapy? How effective is bibliotherapy as a "stand-alone" intervention? How effective is it as an adjunct to other interventions? I hope that this meta-analysis will help

provide useful data to inform these debates. To that end a variety of specific research questions were developed that could be addressed with the available data.

Research Questions

The primary questions that directed the current research were:

1. Is bibliotherapy treatment effective in general (across problem types and dependent variables)?
2. Is bibliotherapy differentially effective across problem types (e.g., is it more effective for helping people overcome phobias than problem drinking? as suggested by Gould and Clum, 1993)?
3. Does the way the dependent variable was measured moderate effect size?
4. Do research methodology variables other than type of dependent variable moderate the results?
5. Does having contact with a therapist during bibliotherapy increase its effectiveness?
6. Does the type of therapist contact (e.g., individual, group, mail) moderate the effects of bibliotherapy?
7. Does the medium of the treatment moderate the effects of bibliotherapy, i.e., are self-help books more or less effective than self-help audio-visual presentations?
8. Are different paradigms of bibliotherapy differentially effective?
9. Are the effects of bibliotherapy moderated by whether the clients' problems are clinical or non-clinical?
10. Does the presence of a cash deposit, fee, or payment have

positive effects for bibliotherapy?

11. Does bibliotherapy have similar effects as therapist administered psychotherapy when they are directly compared? Does adding bibliotherapy to a primarily therapist-directed approach increase the psychotherapy's effectiveness?

12. Is the treatment effect of bibliotherapy maintained past the end of treatment?

The order of the first five research questions is purposeful; they were prioritized according to what previous subjective and meta-analytic reviews of the literature have deemed important. Every review read by this author mentioned problem type as an important moderating variable and several used it as the only or first moderating variable they addressed. Numerous reviews found that bibliotherapy was differentially effective for different problem types, but their assertions (usually based on vote-count methods) about which types were amenable to bibliotherapeutic change did not always agree (see Chapter 2 for details). In the current analysis, it was hypothesized that there would be significant differences between problem types.

Research questions three and four both address potential heterogeneity based on research methodology and publication bias. Research methodology differences (e.g., use of placebo versus no-treatment control groups, different outcome variable types) and availablity bias are frequently cited as potentially important moderating variables in treatment meta-analyses (Lipsey & Wilson,

1994; Durlak & Lipsey, 1991; Glass, McGaw, & Smith, 1981). In the current analysis, it was hypothesized that dependent variable type, control group type, and publication type woulld moderate effect sizes.

Research question five is another that is discussed in several of the literature reviews of bibliotherapy. It was hypothesized in this analysis that having more contact with a therapist would increase the effectiveness of bibliotherapy.

Research questions six to ten are all secondary analyses. They all consist of potentially moderating variables in which adequate data are available to test. However, it was not hypothesized that any of these would actually moderate the effect sizes.

Research questions eleven and twelve are actually unique meta-analyses based on variant datasets. In question eleven, it was hypothesized that therapist-directed therapies that use bibliotherapy as an adjunct would outperform therapist-directed therapies without bibliotherapy which would in turn outperform bibliotherapy that has minimal-therapist contact. In question twelve, it was hypothesized that the effects of bibliotherapy would erode to some degree at follow-up.

CHAPTER II

REVIEW OF THE RELATED LITERATURE

Surveys of the Clinical Use of Bibliotherapy

Several surveys about the current use of bibliotherapy by practitioners have been conducted (Atwater & Smith, 1982; Smith & Burkhalter, 1987; Starker, 1986). According to these surveys, a majority of practitioners recommend bibliotherapy, and they believe it to be helpful to clients. The books recommended appear to be very diverse, but the most common seem to be from the popular press, e.g., When I say no I feel guilty (Smith, 1975), Passages (Sheehy, 1974), The pleasure bond (Masters & Johnson, 1970), What color is your parachute? (Bolles, 1988).

Smith and Burkhalter (1987) reported that 51% of their American Academy of Psychotherapists respondents indicated that they used bibliotherapy in their clinical practice. When asked to rate on a five-point scale how effective they thought bibliotherapy was for particular problems, the following results were reported (from least to most effective): Weight loss, sexual dysfunction, communication skills, marital conflict, and assertiveness. However, Smith and Burkhalter's results must be interpreted cautiously because they had only a 32% response rate to their survey.

Starker (1988) reported that 88.6% of responding psychologists in the Seattle area and 60.3% in the San Diego and Boston areas prescribed self-help books to their clients. Starker also encouraged cautious interpretation of his data, but said he was "reasonably convinced that the prescription of self-help works has become commonplace among practicing psychologists" (p.599).

The results of the present meta-analysis may have important practical ramifications on the findings of these surveys and the use of bibliotherapy by practitioners.

Subjective Reviews of Bibliotherapy's Effectiveness

There have been numerous previous reviews of the literature concerned with bibliotherapy and self-help therapies. To date, however, all but two of these reviews have been of the non-meta-analytic, "vote method" variety.

Glasgow and Rosen (1978), in a subjective review of about 90 studies on therapy manuals, concluded that behavioral bibliotherapy interventions were moderately successful for fear reduction, weight reduction, and study behavior but less successful for smoking cessation and sexual dysfunction. They noted that further research was needed in the areas of child behavior problems and physical fitness before generalizations could be made. They further recommended that future studies be conducted in conditions as similar as possible to the intended use of the materials. They indicated a need for more studies to have as little counselor intervention as possible since that is the most typical use of self-

help literature.

Schrank and Engels (1981), in a "vote method" review of 70 studies, asserted that there is strong evidence for bibliotherapy efficacy in the areas of attitude change (17 of 20 studies reported statistically significant change), psychotherapeutic gains (5 of 5 reported significant results), and assertiveness training (3 of 3 with significant results), but equivocal evidence in the areas of academic achievement, behavioral change, fear reduction, helper effectiveness, marital accord, self-concept improvement and self-development. They also noted that bibliotherapy is an emerging intervention and "that positive recommendations of the value of bibliotherapy exceed available documentation of its usefulness" (Schrank and Engels, 1981, p. 146).

In a short subjective review of about 10 studies, Stevens and Pfost (1982) pointed out that the scientific justification for the use of bibliotherapy had not yet been demonstrated. They also suggested guidelines for articulating future research. These included reporting moderating variables like "(a) type of literature ...., (b) degree of therapist contact ..., (c) client characteristics (age, intelligence, locus of control, etc), (d) clearly defined therapeutic goals ..., (e), duration of bibliotherapy, and (f) use of bibliotherapy alone or as an adjunctive technique" (Stevens & Pfost, 1982, p. 23).

Riordan and Wilson (1989) reviewed approximately 30 studies reported between 1981 and 1988 and found mixed outcomes. They came to the following conclusions: (1) behaviorally based reading

materials have at least some empirical validation; (2) less didactic forms of bibliotherapy (e.g., fiction, poetry) remain essentially unvalidated; and (3) bibliotherapy is of increasing interest to practitioners despite mixed empirical results (Riordan & Wilson, 1989).

Perhaps the most comprehensive review of the field to date was done by Craighead, McNamara, and Horan (1984). In their review of 92 behaviorally-based studies, they indicate that totally self-administered programs do appear effective for "particular individuals, but that a majority of people seem to want or need some therapeutic contact" (p. 920). They further cited that minimal-contact (with therapists) bibliotherapy did appear to be quite cost effective, but that therapist-administered conditions (more contact) seemed to be slightly superior.

They made several assertions about the differential effectiveness of bibliotherapy across problem types. They proposed that self-help treatment of problem-drinking might have as much success as traditional modes of therapy. They pointed out that the procedures used in these problem-drinking studies might not be generalizable to “alcoholic” subjects. For obesity and smoking they thought that totally self-administered programs were not particularly effective. They found more support for the effectiveness of bibliotherapy on assertion, depression, anxiety, vocational concerns, sexual problems, and academic problems.

Craighead, et al., also provided an extended discussion about the

possible positive and negative effects of bibliotherapy. For example, they pointed out that bibliotherapy may be very suitable and effective for mild cases of depression, but there "are serious ethical considerations with the more severely depressed" (1984, p. 920). Other possible negative effects of self-help orientations included the promotion of non-problems and inappropriate self-diagnosis. The promotion of non-problems means that people dealing with normal developmental issues may label them as mental problems. Inappropriate self-diagnosis may lead a person to underestimate the severity of their problem and their need for more intensive assistance. Another warning they issued was that future studies need to determine the degree to which client characteristics, severity of the problems,and need for contact with a therapist/helper may moderate the effectiveness of bibliotherapy.

Meta-analyses of Bibliotherapy

Scogin, Bynum, Stephens, and Calhoon (1990) conducted a meta-analysis of 40 bibliotherapy studies. They found effect sizes (d) of 0.96 for self-administered treatments over controls and 1.19 for bibliotherapies with some minimal therapist contact over controls. They found nonsignificant differences (effect size approximately 0.10) between self-administered and therapist-administered treatments. There was some evidence that combined self- and therapist-administered interventions were more effective than self-administered bibliotherapy alone. They failed to report any findings on homogeneity of effect sizes. While their

findings did support the effectiveness of self-administered programs, the authors pointed out that "the majority of studies dealt with rather circumscribed problems that may lend themselves to more education- and information-based interventions" (p. 45) and that the materials evaluated by researchers are rarely those that are prescribed by practicing psychologists.

Their study was also limited by the relatively low number of studies and by questionable grouping strategies. For example Scogin, et al., collapsed smoking and weight loss studies together under the title of habit control. However, Craighead, et al. (1984) reported that bibliotherapy appeared more effective for weight loss than it did for smoking cessation.

Gould and Clum (1993) conducted a meta-analysis on 40 studies of self-help treatment approaches and found a mean effect size (d) of 0.76 at posttreatment and 0.53 at follow-up. In addition they found that studies using placebo control groups had smaller effect sizes than those using no treatment control groups and that some problems (fears, depression, headaches, and sleep disturbances) were more amenable to self-help than habit disturbances (smoking, drinking, overeating).

These two meta-analyses, however, have several weaknesses. Both used only published studies in their database, potentially allowing for a "publication bias" (Light and Pillemer, 1984). Both found a relatively small numbers of studies (40). Although both these meta-analyses had found 40 studies that met their similar

inclusion criteria, they shared only 9 studies between them. This suggests that their databases may not have been representative and their searches not entirely adequate. Perhaps most important, both used the arguably outdated meta-analytic technology of Glass, McGaw and Smith (1981) rather than the more empirically defensible technologies of Hedges and Olkin (1985) or Hunter and Schmidt (1990).

Meta-analytic Reviews of Psychotherapy

Lipsey and Wilson (1994) quantitatively reviewed 302 meta-analyses, of which 90 were meta-analyses of psychotherapy, counseling, or psychoeducational interventions. The mean unweighted effect size for these 90 meta-analyses was .588 and .571 when weighted by number of samples in each meta-analysis (number of samples in each ranged from six to 475).

A brief review and critique of the seminal meta-analysis of psychotherapy, Smith and Glass (1977), may help illuminate some factors that could be important in the current meta-analysis. Smith and Glass meta-analyzed 375 controlled evaluations of psychotherapy and counseling. They found an overall effect size (delta) equal to 0.68. When broken down by problem type (what they called "type of outcome measure, p. 756), they found fear-anxiety reduction (E.S.=.97) and self-esteem (E.S.=0.90) to be "more susceptible to change in therapy" than the categories of "adjustment" and "school/work achievement" (1977, p. 756). The adjustment category included more serious behavior manifestations

like psychosis, alcoholism, criminal episodes, and depression.

Smith and Glass also analyzed for type of therapy as defined by theoretical viewpoint. When they compared the two "superclasses" of behavioral and nonbehavioral strategies with data drawn only from 119 studies in which they were simultaneously compared with the same control, they found only marginally different (nonsignificant) effect sizes of 0.69 and 0.62, respectively (p. 758). When comparing many different types of therapies (e.g., systematic desentization, Rational-Emotive, client-centered), they asserted that "few important differences in effectiveness could be established among [them]" (1977, p. 752).

Brown (1987) reviewed six meta-analyses of psychotherapy, including the Smith and Glass (1977) study and several that used subsets of their database. He criticized several of these meta-analyses for nonindependence, that is some studies contributed more than one effect size to the analysis. However, he admitted that regardless of whether the meta-analyses used independent or nonindependent effect sizes, the mean effect sizes ($\underline{d}_{++}$) were always between 0.68 and 0.93, suggesting that psychotherapy was moderately effective. He also pointed out that variables like type of control group, subject IQ, and outcome measure reactivity seemed to moderate the effects of psychotherapy in these meta-analyses. He criticized the meta-analyses for publication bias and the failure to report "Fail-Safe N" statistics.

Matt (1989) criticized the Smith and Glass meta-analysis for

several of its decision rules, most notably the conceptual redundancy rule. This "rule aims to exclude effect sizes based on any outcome measures that are redundant...." (Matt, 1989, p. 107). In other words, Smith and Glass (1977) did not include all effect sizes computable from their data; they selectively excluded those that were judged redundant in magnitude and outcome type. Matt questioned, then, whether this rule could be implemented reliably and had independent coders compute effect sizes from a random sample of the studies used by Smith and Glass. These independent coders computed effect sizes that were typically half the magnitude that Smith and Glass had found.

Summary

In summary, there is moderate amount of evidence that bibliotherapy is somewhat effective for some people with some types of problems. In some cases its effectiveness may even be comparable to traditional psychotherapy. However, despite the facts that most clinicians do prescribe bibliotherapy to clients and that many non-clients bibliotherapeutically treat themselves, the evidence for its effectiveness has not been systematically analyzed. We do not yet know for whom and under what conditions bibliotherapy might be most effective.

CHAPTER III

METHOD

Definition of Bibliotherapy Treatments

The definition of bibliotherapy that will be used in this study will be "the use of written materials or computer programs, or the listening/viewing of audio/videotapes for the purpose of gaining understanding or solving problems relevant to a person's developmental or therapeutic needs. The goals of the bibliotherapy should be relevant to the fields of counseling and clinical psychology." This is a definition most similar to those used by Schrank and Engels (1981), Smith and Burkhalter (1987), and Craighead, McNamara, and Horan (1984). Bibliotherapy is perhaps most frequently thought of as reading self-help books to solve personal problems. The current definition includes this activity, but also expands the definition to include self-help audio-visual therapies (e.g., using computer programs or watching videotapes). As suggested by Craighead, et al., "the technological revolution has redefined the book as only one of several possible modes" (1984, p. 878). This definition also deems important self-help for developmental needs as well as therapeutic problems. Much of the bibliotherapy available today is directed towards helping with normal developmental needs and it seems theoretically plausible

that bibliotherapy could be more beneficial (and ethically less troublesome for clinicians) for "normal" persons with adequate ego strength and self-confidence. The current definition does not include use of self-help groups like Alcoholics Anonymous, only self-help provided through some medium like books or computers.

To be included in this review, a study had to meet the following criteria: (a) The treatment must correspond with the definition of bibliotherapy in the previous paragraph; (b) it must have included adults only (college aged persons included) working on their own concerns and not those of their children (e.g., parents learning to better discipline their children would not be included); (c) it must not be a media-based campaigns (e.g., Stanford's heart disease studies); (d) the bibliotherapy must have been compared with a comparison group drawn from the same population as the treated subjects; (e) the bibliotherapy must be a part of primary treatment strategy rather than a post-treatment maintenance strategy; (f) the bibliotherapy must be described as longer than 1-2 pages (e.g., a booklet or a series of handouts); (g) the study must have been reported in English; and (h) the data reported must be amenable to meta-analytic procedures. In addition, studies (dissertations) that were only available through University Microfilms Inc. (UMI) were not included because of their prohibitive expense (approximately $38 each). However many studies available from UMI were garnered via other sources. Media-based campaigns, while conceptually similar to bibliotherapy, were not included in the

current meta-analysis because their reports are generally unable to indicate how many of their subjects actually used the "therapeutic procedures."

The search was not limited to particular years of study publication; articles as early as the 1940's were investigated for possible inclusion into the database. However, no studies reported prior to 1968 were eventually included in the final database because none met the inclusionary definitions stated above.

Some bibliotherapy studies include comparisons between traditional therapist-administered psychotherapies, totally self-administered bibliotherapies, and therapies which combine bibliotherapy with therapist contact. In this study there was no attempt to break these into three separate categories (like Craighead, et al., 1984), only two, bibliotherapy and therapist-administered. For the current meta-analysis, the traditional therapist-administered psychotherapies were coded as a distinct category, but if a research report indicated that the therapeutic mode included bibliotherapy, it was included in a single bibliotherapy category and the amount of therapist contact (in minutes and weeks) was coded for further analysis. In a few cases, studies used in this meta-analysis did not have a control group, but indirect effect size comparisons could be made between different bibliotherapies and therapist-administered psychotherapies within the same study.

Literature Search

Several methods were used to ensure comprehensiveness. First, computer searches were conducted on PsychLit, Dissertation Abstracts, Educational Resources Information Center, and Infotrac, using "Bibliotherapy" as the key word. Checks of other keywords (e.g., "self-help," "self-administered") provided no additional useful citations. Second, searches of the bibliographies of the review articles cited in Chapter 2 (Glasgow and Rosen,1978; Schrank and Engels,1981; Stevens and Pfost, 1982; Riordan and Wilson,1989; Craighead, McNamara, and Horan, 1984; Scogin, Bynum, Stephens, and Calhoon,1990) were also completed. These provided the largest numbers of studies (N= 203).

Third, further searches were also conducted on the bibliographies of books about bibliotherapy (e.g., Rubin, 1978) and the reference lists of other articles already included in the data pool. Finally, hand searches of five journals that frequently had reports about bibliotherapy were also conducted: Behavior Therapy, Behavior Research and Therapy, Journal of Counseling and Development, The Journal of Poetry Therapy, and Journal of Counseling Psychology. These journals were searched issue by issue for the years 1970-1992. When a limited search prior to 1970 was conducted on these journals, no pertinent studies were found. Using all the methods described above, only two pertinent studies prior to 1970 were found, none prior to 1968.

The initial data base generated through these processes

numbered over 400. Articles that were obviously not empirical studies (e.g., "A bibliography of books for children's bibliotherapy;" "A librarian's perspective on bibliotherapy") were excluded from the initial database. The resulting database of potential inclusions numbered 276. Of these, 81 were unpublished and 195 were published. Of the 81 unpublished studies, 49 were unavailable because of prohibitive costs (i.e., available only through University Microfilms, Inc.). Of the published studies, 11 were reviews of the literature (of bibliotherapy or a related topic like smoking reduction research), 43 were articles or studies of children's bibliotherapy, ten were opinion articles, three were qualitative evaluations, and one was in a foreign language not read by this researcher.

The remaining 159 articles (127 published and 32 unpublished) included 21 investigations that could not be used because they did not employ a control group for comparisons. Twenty-four studies provided insufficient data from which to compute effect sizes (e.g., means with no standard deviations or inferential statistics, factorial anova results without summary tables). Thirty-five were judged to not be a study that met the current meta-analysis's definition of bibliotherapy (e.g., self-help strategies taught to subjects without the use of bibliotherapy, bibliotherapy used only as a post-treatment strategy, media-based campaigns).

The result was a database of 79 useful study samples. The references for this final database are listed in Appendix H. Nine of these 79 did not employ a control group, but did directly compare a

therapist-directed treatment with a bibliotherapy treatment group. It was, therefore, possible to compute an effect size comparison with these nine studies, but only include them in the analysis comparing therapist-directed treatments to bibliotherapy. The result was a final database of 70 useful samples in the main meta-analysis, of which nine were unpublished. The unpublished studies represented 13% of the database for the main analysis. In the secondary analysis comparing therapist-directed treatments to bibliotherapy, 30 samples were used, of which five (17%) were unpublished studies.

Classification and Coding Systems

The following variables were coded in this study:

1. Design characteristics (group assignment methods, presence and type of control group, statistical analysis used to determine effect size);

2. Publication type (journal publication, dissertation, or other);

3. Amount of therapist contact (in average minutes per week and frequency of sessions) and length of treatment;

4. Type of therapist contact (None; individual face-to-face; group meetings; phone contacts; weigh-ins);

5. Measurement of dependent variable (validated self-report scale, non-validated self-report scale, observed behavior, self-reported behavior, academic achievement, scale rating by another person, physiological measurement);

6. Type of problem. There were 13 different problem types that were originally coded. The types coded include problems with Alcohol, Anxiety, Assertiveness, Career Indecision, Depression, Self-Esteem/Self-Concept concerns, Habit control (other than smoking and eating), Marital/couple dissatisfaction, Sexual dysfunction, Smoking, Studying problems, Test Anxiety, and Weight loss. If the problem was not one of these, it was coded Other.

Of those 13 problem types, ten had sufficient numbers of studies ($n > 3$) to allow the possibility of intergroup comparisons. The vast majority of bibliotherapy studies (86%) addressed one of these 10 problems. The problem types of Alcohol, Habit Control, and Marital/Couple Satisfaction had insufficient numbers for intergroup comparisons.

In a few cases, categories originally coded separately were later combined because they were small in number, conceptually similar and/or statistically homogeneous. For example, in some analyses Alcohol, Habit Control, and Smoking were combined into an Impulse Control category. Weight loss was never included within this category. Anxiety and Test Anxiety were eventually combined because they were statistically homogeneous and conceptually similar. In some analyses Marital/Couple satisfaction was combined with the Other category because of insufficient numbers of studies.

7. Type of reading material (Manual or general publication);

8. Length of material (number of pages or audio/video minutes)

9. Mean or median age of subjects (mean used if both cited);

10. Gender ratio of sample;

11. Education level of subjects;

12. Reading ability of subjects and the readability of the bibliotherapy;

13. Personality style match (a few studies looked at differential effects with subjects of various personality styles, e.g., locus of control, attribution style, Holland code, Myers-Briggs code);

14. Medium of bibliotherapy (paper, audiotape, videotape, informational computer, interactive computer);

15. Type of reading matter (direct or indirect instruction; indirect instruction includes readings of a fictional, poetic, or metaphorical nature);

16. Compliance to reading program;

17. Drop-out rate;

18. Presence & amount of cash deposit or other incentives;

19. Sample size;

20. Psychological paradigm (e.g., behavioral, cognitive, humanistic, other, cognitive-behavioral, unspecified);

21. Length of treatment in weeks;

22. Time period between treatment and evaluation;

23. Source of subjects (college students solicited from courses, college students solicited from the general campus community, subjects solicited from the general community, referrals garnered from mental health treatment facilities, or other. The only subject samples coded "other" both came from prison

populations);

24. Extra-biblio homework (the presence of homework assignments as part of the bibliotherapy, e.g., journaling, trying-out new behaviors);

25. Training level of therapist/contact person (paraprofessional, Master's level counselor, Ph.D. psychologist);

26. Clinical population or not;

27. Severity of clinical problem (estimated by finding norms of dependent variables);

28. How effect size was computed (e.g., post-test scores vs. gain scores).

For more specifics about these coding processes, see Appendix A for examples of the coding sheets and Appendix B for the coding sheet training manual.

Interrater Agreement and Reliability

Each study was coded for study characteristics and effects sizes by the primary investigator. Twenty-seven studies (34%) were independently coded for study characteristics by a graduate student in education (Judge A). Twenty-six studies (33%) were independently coded for effect sizes and variable types by a doctoral candidate in counseling psychology knowledgeable of meta-analytic procedures (Judge B). A total of 127 effect sizes were computed from these 26 studies by Judge B as part of the reliability check.

Categorical variables that were possible moderators were analyzed for interrater agreement using Cohen's kappa (Tinsley and

Weiss, 1975) and absolute agreement. Continuous variables were analyzed for interrater reliability using the intraclass correlation coefficient (Tinsley and Weiss, 1975) and the Pearson product-moment correlation coefficient. Since between judges variance in this application should be considered as error, the intraclass correlation was computed using a standard one-way ANOVA procedure. This type of intraclass correlation also allows for generalizablity to other potential judges. The results of these reliability checks are summarized in Table 1.

Table 1

Interrater Agreement and Reliability Estimates

Categorical Variables	Kappa	Absolute Agreement
Group Assignment Process (e.g., Random, Matched)	.842	.963
Control Group Type (e.g, Placebo, No treatment)	.948	.963
Therapist-only Treatment Comparison	.924	.963
Publication Type	1.000	1.000
Type of Therapist Contact (e.g., Individual, Group)	.768	.815
Therapist Training Level	.842	.889
Was Researcher a Therapist?	.879	.926
Problem Type (e.g, anxiety,assertion)	1.000	1.000
Biblio Either a Manual/Publication	.625	.852
Biblio Treatment Medium (e.g., Paper, Computer)	1.000	1.000
Instruction Type (i.e., Didactic or Affective)	1.000	1.000
Presence of Cash Deposit	1.000	1.000
Source of Subjects (e.g., college class sample)	.641	.741
Treatment Paradigm (e.g., Behavioral)	.945	.963
Presence of Homework Assgnt	.855	.963

Table 1 continued

Categorical Variables	Kappa	Absolute Agreement
Dependent Variable Type (e.g., physiological, self-report)[a]	.982	.969
Mean Agreement - Categorical	.891	.938
Median Agreement - Categorical	.935	.963

Continuous Variables	Pearson	Intraclass Corr.
Amount Therapist Contact	.724	.690
Frequency of Contact	.916	.905
Length of Treatment	1.000	1.000
Sample Size	.989	.988
Number of Biblio Pages	1.000	1.000
Age of Clients	1.000	1.000
Client Gender (% Female)	1.000	1.000
Education Level of Client	1.000	1.000
% Clients with Post-Secondary Ed.	1.000	1.000
Mortality Rate of Study	.840	.693
Effect Sizes (k=127) [a]	.960	.986
Mean Reliability - Continuous	.948	.933
Median Reliability - Continuous	1.000	1.000

Note. [a] denotes coding by Judge B. All others were coded by Judge A.

The two categorical variables with the lowest interrater agreements (kappa < .650) were due to systematic problems in the training of Judge A. For the source of subjects variable, Judge A coded studies as "mental health referral" very rarely because she confused it with "solicited non-college population." For "Type of Bibliotherapy Material" Judge A forgot to check the references section on five studies to determine if the researchers had used an unpublished manual or a published book.

All disagreements between judges were resolved through discussion.

In addition to the reliability and agreement of coded variables, there was the potential for disagreement as to whether particular samples within studies should be coded as Therapist-administered groups, Bibliotherapy groups, Control groups, or not included at all. For example, there were several studies that gave a bibliotherapy treatment to a particular sample, but called that group a "placebo control" group. There were other studies that gave a bibliotherapy treatment, but not one that some coders might think was a treatment for that problem type (e.g., a Logotherapy book for the treatment of depression). Both Judges A and B coded for type of group on different studies. The absolute agreement between Coder A and the primary investigator for placement of all groups within a study was .852 on 27 studies. The absolute agreement between Coder B and the primary investigator was .731 for 26 studies. The overall agreement was .792 ($\underline{k}$=52). These numbers are a

conservative estimate of agreement since within the studies where there was some disagreement, there was typically more agreement than disagreement. For example, the two raters might have agreed on the placement of four groups in a study that had only five, but the one disagreement would have resulted in coding this study as a "disagreement." Cohen's kappa was not computed because it requires a known and limited set of categories from which to compute chance agreements. With this "variable" each study differed in its complexity and chance agreement. As before, disagreements between judges were resolved through discussion.

In general, interrater agreement and reliability appeared adequate.

Computation and Analysis of Effect Sizes

The effect size used in the main analysis and all subsequent analyses was $\underline{d}_{++}$, the unbiased effect size described by Hedges and Olkin (1985). To calculate $\underline{d}_{++}$,an effect size estimate ($\underline{d}$) was first calculated for each sample by subtracting the comparison group mean from the experimental group mean and dividing by the pooled standard deviation. Then, because d has been shown to be a biased (overestimated) index of the population effect size in small samples (Hedges, 1981), each effect size was weighted by the inverse of its estimated variance (to give larger weight to studies with smaller variances). These unbiased effect sizes (Rosenthal & Rubin, 1982; Hedges & Olkin, 1985) were computed for each outcome variable of each study using the DSTAT software program (Johnson, 1989)

The following procedures were used in calculating effect sizes. If means and standard deviations were made available by the study author, these were used to compute effect sizes even if other statistics (e.g., t-tests, one-way F's) were provided. If more than just one treatment and one control group were included in a study (e.g., bibliotherapy with 40 minutes of counselor contact, bibliotherapy with no counselor contact, no-treatment control, and placebo control in one study), the standard deviations of all pertinent groups were pooled into one standard deviation for analysis for each group. The DSTAT program allows this procedure to be done efficiently. This strategy was used because it was believed that this overall pooled standard deviation would generally be a better estimate of the population standard deviation. A few internal Monte Carlo checks exhibited little difference between effect sizes computed with an overall pooled standard deviation versus those in which only two group standard deviations were used.

If there happened to be multiple treatment groups, but these did not differ on attributes coded for in this meta-analysis (e.g., behavioral studies that differentiated between self-reward and self-monitoring groups), those groups were combined together. Pooled means and standard deviations were weighted for group size. In two cases, standard deviations were not provided by the study author, but the raw data were and I was able to compute standard deviations from them.

Effect sizes were computed only on post-test scores unless the study author only provided data on change scores. In cases where pretest scores were reported and it was apparent that the pretest mean differences between treatment and control groups were greater than .50 standard deviations, these were not included in the analysis unless the data were also provided as change scores (with standard deviations) or adjusted by use of ANCOVA. To include post-test effect sizes that differed by too much at pretest confounded the results. No samples were excluded from the meta-analysis because of this rule, but a few outcome variables were ignored and a few effect sizes defaulted at zero. For example, Glasgow, Swaney, and Schafer (1981) reported pretest and posttest means and standard deviations along with a nonsignficant ANCOVA. The posttest mean differences in habit control (nail-biting) between treatment and placebo-control groups suggested an effect size of 0.22. However, the pretreatment mean differences of some groups differed by nearly one standard deviation (e.g., nail lengths of 1.08 and .93 centimeters with standard deviations of .17 and .20, respectively. Because of these pretreatment differences it was thought safer and more defensible to use the default of $\underline{d}$=0.00 from the nonsignficant ANCOVA results. This decision was further supported by nearly identical change scores across groups. Application of this rule was necessary in only four samples.

If means and standard deviations were not available, but useful inferential parametric statistics were provided (e.g, t-tests,

F-tests), then these were used to compute effect sizes. The DSTAT program transforms these inferential statistics into d's by finding pooled standard deviations and standard differences between group means. Johnson (1989) reported the specific equations used in Appendix A of the DSTAT program manual and cited from where these equations were taken (e.g., Glass, McGaw, & Smith, 1981; Winer, 1971). If both ANOVA and ANCOVA statistics were reported in a study, the ANCOVA results were used to compute effect sizes.

If the data were frequencies or proportions, DSTAT treats each proportion as the mean of a distribution of 0's and 1's and a pooled standard deviation is derived from a binomial statistic (Johnson, 1989, p. 105). An effect size is then computed by subtracting those "means" and dividing by the resultant pooled standard deviation. If the frequencies or proportions were not reported, but a chi-square was, this was used to compute the effect size. DSTAT converts chi-squares into Pearson correlations and then converts those into effect sizes (Johnson, 1989, p. 104-105).

As a last resort, if only probability levels were reported, these were used to estimate an effect size. DSTAT computes this effect size by converting the p-value to a z score, transforming the z to an r and then transforming the r to an appropriate effect size (d). If a study simply reported that there were no significant differences between groups, the default effect size became zero. This default rule of zero was necessary only with two samples in the main analysis.

The only outcome measures that were not used to compute effect sizes were those that did not seem to be relevant to the treatment procedures. For example, a few studies used locus-of-control as a secondary outcome measure for treatment of depression or weight loss treatment. In these cases only the primary outcome measures (e.g., the Beck Depression Inventory or number of pounds lost) were used to compute effect sizes.

There were seven different types of dependent variables used by studies in this meta-analysis: Physiological measures (e.g., weight loss, biofeedback), behavior observed by the researcher(s), scale rating by others (e.g., non-behavioral ratings such as husband's estimate of wife's mood change), standardized self-reported measures (e.g., the MMPI, the Beck Depression Inventory), unstandardized self-report measures, academic achievement (e.g., GPA), and self-reported behavior (e.g., cigarettes smoked in past week, estimated minutes to ejaculation).

A specific strategy was implemented for combining effect sizes within studies that had multiple outcome measures. First effect sizes were computed on each individual outcome measures. For example, a bibliotherapy study on test anxiety might operationalize that construct several ways. The researchers might measure the outcome via two biofeedback readings (EMG and GSR), three separate standardized self-report measures of test anxiety, one self-estimate on an "anxiety thermometer", and the students GPA improvement after bibliotherapy completion.

Then, these individual effect sizes were grouped and averaged within the study according to their dependent variable type. For example, Instead of combining all seven of the aforementioned test anxiety measures together for an mean effect size, a mean effect size was first computed for each of the four dependent variable types (physiological, standardized self-report, non-standardized self-report, and academic achievement). These effect sizes were kept separate in one data base to allow a moderator analysis of dependent variable type. However, for the main analysis, if a study had several types of dependent variables, these were then averaged to obtain an overall effect size for that study. For the main analysis each study contributed only one effect size. This strategy was a judgment call (Wanous, Sullivan, & Malinak, 1989) similar to recommendations made by Durlak and Lipsey (1991).

A similar strategy was used with studies that had both placebo and no-treatment control groups. These effect sizes were left separate for a moderator analysis, but were combined for the calculation of the overall effect size estimate.

To arrive at a single estimate of the overall effect size ($\underline{d}_{++}$) in each meta-analysis, study outcomes (individual $\underline{d}$'s) were then combined by finding a weighted average of the $\underline{d}$'s using a procedure described by Hedges and Olkin (1985). This procedure involves multiplying each effect size by its weight (i.e., by the inverse of its variance) and dividing the sum of these multiplications by the sum of the weights. After this, 95% confidence intervals around $\underline{d}_{++}$

were calculated using standard procedures.

To determine whether samples in each meta-analysis shared a common effect size (i.e., were homogeneous), the unbiased mean effect sizes were tested for homogeneity by the Q_T statistic (Hedges & Olkin,1985). Q_T is distributed as a χ^2 with k-1 degrees of freedom (where k = total number of studies). In cases where Q_T was significant, possible moderating variables were studied. For categorical variables a grouping strategy based on the research questions listed in Chapter I was used. Then the Q_B statistic is computed to test for between group differences. The Q_{wi} statistic is computed to determine if there is significant within-class variability. The Q_B and Q_{wi} are evaluated together. A model is considered well specified if Q_B is significant, but Q_{wi}'s for the resulting categories are nonsignificant. If Q_{wi} remains significant, the hypothesis of whether a moderator variable accounted for the original heterogeneity is insufficient (Durlak & Lipsey, 1991).

After this, 95% confidence intervals around each d_+ (the symbol for mean effect size within a category) can then be used to determine if particular categories are significantly different from others within an analysis.

For continuous variables, Rosenthal and Rubin's focused comparison method (1982) was used. This method weights studies by sample size and analyzes whether continuous variables are significant predictors of effect sizes. Weighting studies by sample size is recommended in meta-analysis so larger studies will

contribute more to the relationship than smaller studies. In Monte Carlo studies, this focused comparison method reportedly yields conclusions highly similar to the Hedges and Olkin style weighted regression (Johnson, Mullen, & Salas, 1993). The results are reported as an inferential $\underline{z}$-test, not as a magnitude statistic. Positive $\underline{z}$ results indicate a positive linear relationship, negative $\underline{z}$ results indicate a negative linear relationship.

When the categorical and/or continuous model analyses fail to account for significant heterogeneity, the DSTAT program allows for the systematic inspection and removal of outliers cases. The program identifies how much of the $\underline{Q}_{wi}$ statistic each sample accounts for and then allows the meta-analyst to remove those samples that would result in the greatest reduction in heterogeneity. A rule-of-thumb target of five percent or fewer studies removed as outliers will be implemented in the current meta-analysis.

In cases where it was possible, a direct within-study comparison was made because this can help to rule out confounds between effect sizes and other study characteristics (Shadish and Sweeney,1991). This was then done by making a direct comparison between No Treatment versus Placebo/Comparison groups and between Therapist-only versus Bibliotherapy within the studies that had both.

Finally, a failsafe N (Orwin, 1983) was also calculated; this procedure estimates how many additional studies averaging null

results would be needed to reduce the $\underline{d}_{++}$ results to a neglible level. For the current meta-analysis, a neglible level will be defined as the maximum $\underline{d}_{++}$ that could still cause the 95% confidence interval to span zero.

CHAPTER IV

RESULTS

Study Characteristics

There were a total 4677 subjects in the 79 samples. Of these, 2315 received a bibliotherapy, 455 received a therapist-directed therapy (without bibliotherapy), and 1907 were in control groups. Seventy studies had a control group; nine only compared a therapist-directed treatment against a bibliotherapy treatment.

Table 2 summarizes the continuous variables coded. In some cases the k's are less than 79 because not all studies reported on that characteristic. In other case the k's are greater than 79 because there were multiple groups reported within a study.

In general, the studies averaged 59 subjects and retained about 87% of them through post-treatment. The participants averaged in their mid 30's, were well-educated, included more women than men, and met with a therapist a mean of 36 minutes per week,.

Table 3 summarizes the categorical variables coded. In general the bulk of the studies employed randomly assigned comparison groups (84%), used a book as the treatment medium (80%), used direct instruction (96%) rather than indirect/affective approaches, and did not use samples from clinical populations (74%).

Table 2

Characteristics of Bibliotherapy Studies - Continuous Variables

Study Characteristic	M	SD	Range	k [a]
No. of Clients in Sample	56.63	41.03	10-247	79
Clients per Biblio Group	24.00	20.47	5-123	98
Percentage Subjects Retained to Posttreatment	88.37	12.88	31.3-100	68
Percentage Female Clients	64.43	29.86	0-100	71
Client Age	35.31	12.90	18.6-70.5	45
Weeks of Treatment	6.36	3.56	1-15	66
Minutes of Weekly Therapist Contact	38.23	53.10	0-240	74
Length of Bibliotherapy (in pages or minutes)	212.00	211.24	23-960	29
Education Level of Subject (yrs)	13.31[b]	0.83	12-17.1	47
% with Some Post-secondary Ed.	97.36	9.72	41-100	44
Amount of Deposit/Fee	$29.50	27.69	$4-125[c]	22

Note. [a] In this and subsequent tables, k denotes the number of studies. [b] Studies that used a college population but did not specify average grade level were set at 13.1 as a default. [c] Most studies did not report having a deposit, fee, or payment. These numbers are based only on those 22 studies that reported such a financial transaction.

Table 3

Characteristics of Bibliotherapy Studies - Categorical Variables

Study Characteristic	k	%
Group Assignment Method		
Random	66	83.5
Matched	3	3.8
Self-selected	4	5.1
Other	2	2.5
Unreported	4	5.1
Control Group Type [a]		
No treatment	51	64.6
Placebo/Comparison	29	36.7
None (Pre-Post only)	9	11.4
Publication Type		
Journal	68	86.1
Dissertation/Thesis	11	13.9
Training Level of Therapist/Contact persons		
Paraprofessional	8	10.1
Graduate Student	27	34.2
Masters Degree	2	2.5
Mixed Levels	11	13.9
Unreported	31	39.2

Table 3 continued

Study Characteristic	k	%
Dependent Variable Type [b]		
Physiological Measure	14	20.0
Observed Behavior	20	28.6
Scale Rating (By other person)	5	7.1
Self-rate/Standardized	41	58.6
Self-rate/Nonstandardized	15	21.4
Academic Achievement	9	12.9
Self-Reported Behavior	14	20.0
Problem Type [c]		
Alcohol	3	3.8
Anxiety	12	15.2
Assertiveness	12	15.2
Career Indecision	4	5.1
Depression	5	6.3
Habit Control	1	1.3
Marital/Couple	1	1.3
Self-Esteem/Concept [d]	4	5.1
Sexual Dysfunction	5	6.3
Smoking	5	6.3
Studying	5	6.3
Test Anxiety	4	5.1
Weight Loss	11	13.9
Other [d]	7	8.9

Table 3 continued

Study Characteristic	k	%
Type of Reading (or A-V)material		
Manual	48	60.8
General Publication	30	38.0
Unreported	1	1.3
Treatment Medium		
Paper (e.g., book)	63	79.7
Audio or Video	13	16.5
Computer	2	2.5
Several Medium	1	1.3
Instruction Type		
Direct Instruction	76	96.2
Indirect/Affective	1	1.3
Mixed	2	2.5
Presence of Cash Deposit, Fee, or Payment		
Yes	23	29.1
No or Unreported	56	70.9
Subjects from a Clinical Population?		
Yes	21	26.6
No	54	68.4
V code	4	5.1

Table 3 continued

Study Characteristic	k	%
General Paradigm of Bibliotherapy		
Behavioral	29	36.7
Cognitive	10	12.7
Cognitive-Behavioral	24	30.4
Mixed	2	2.5
Unspecified	14	17.7
Did the Bibliotherapy Require Homework Assignments Other than the Required Reading?		
Yes	43	54.4
No	3	3.8
Default	33	41.8
Year of Study [e]		
Pre1965	0	0.0
1965-69	2	2.5
1970-74	19	24.1
1975-79	25	31.6
1980-84	22	27.8
1985-89	9	11.4
1990-92	2	2.5

Table 3 continued

Study Characteristic	k	%
Source of subjects		
Solicited from a College Classroom	22	27.8
Solicited from a College Population	16	20.3
Solicited from Non-college Population	31	39.2
Traditional Referral Clients	8	10.1
Other (Prison)	2	2.5
Studies Reporting Reading Ability of Subjects	5	6.3
Studies Reporting Readability of Material	3	3.8
Studies Reporting Compliance to Therapy	36	45.6
Studies Reporting Pretreatment Matching of Clients to Therapy by Personality Type	4	5.1
Studies Reporting Years of Experience of Therapists	10	12.7
Studies with Follow-up Reports	29	41.4 [f]

Note. The total percentages in this table were computed from the total collection of samples (k=79) unless otherwise noted below.

[a] The Control Group numbers do not total to k=79 or 100% because 10 studies had both a No treatment and a Placebo control group.

[b] The number and percentages for Dependent Variable Type are based only on the bibliotherapy samples from the the 70 studies with control groups. The percentages do not total to 100% because many

studies used more than one dependent variable type. [c] The Problem type data presented above is based on the entire sample of 79 samples. Of the nine studies without control groups that were not included in the main analysis, two were on Weight Loss, two on smoking, and one each on Alcohol, Anxiety, General Counseling, Sexual Dysfunction, and Other. [d] A more complete description of the studies included in the Self-concept/Esteem and Other categories is available in Table 4. [e]The mean and median publication dates were 1978. [f] This percentage is based on the samples only in the main analysis ($\underline{k}$=70) because no follow-up analysis was conducted on the therapist-to-bibliotherapy comparison.

As can be seen at the end of Table 3, neither the reading ability of subjects, the readability of the material used, nor personality type as a potential moderating variable were frequently reported.

Unbiased Effect Size Estimate and Homogeneity of Effect Sizes - Evidence for the General Effectiveness of Bibliotherapy

The summary data presented in the first row of Table 4 provide an overall estimate of the effect size across the 70 studies included in this meta-analysis. This estimate condenses studies regardless of problem type, amount of therapist contact, type of control group, and type of dependent variable. The unbiased effect size estimate ($\underline{d}_{++}$) was .565. Its 95% confidence interval of .494 to .636 did not span zero. Thus, the null hypothesis of a zero effect size could be rejected. Appendix C provides a graphic display of the 70 effect sizes.

The overall effect size estimate of .565 suggests a moderate degree of bibliotherapy effectiveness. However, the calculation of the homogeneity statistic, $\underline{Q}_T$ = 224.543, $\underline{p}$=.000 (see also first row, Table 4) indicated significant heterogeneity among the effect size estimates.

Tests of Categorical Models - Problem Type as Moderator

The second research question of this meta-analysis was whether significant heterogeneity obtained among effect size estimates might be explained by types of problems treated. The data presented in Table 4 suggest that significant between groups

heterogeneity did exist (Q_B=53.483, p=.000) and that the heterogeneity within the groups (Q_{wi}) was explainable by chance for seven of the ten problem types. Problem types are ranked from highest to lowest effect sizes in this and subsequent tables.

Table 4

Bibliotherapy Effect Size Estimates and Tests of Problem Type as a Moderator Variable Across All Types of Dependent Variables

Sample/Category	k	d_{++}	d_{+}	95% CI	Q_B	Q_T	Q_{wi}	p
Total Sample	70	0.565		0.494/0.636		224.543		.000
Problem Type							53.484	.000
Sexual Dysfunction	4		1.279	0.794/1.863			6.166	.187
Assertion	12		0.946	0.735/1.158			33.839	.001
Anxiety [a]	15		0.906	0.731/1.080			39.153	.001
Depression	5		0.567	0.246/0.887			9.358	.096
Career	4		0.537	0.314/0.760			9.403	.052
Others [b]	7		0.524	0.305/0.743			12.893	.075
Self-esteem/concpt [c]	3		0.515	0.133/0.896			6.322	.097
Weight Loss	9		0.396	0.215/0.576			41.921	.000
Studying	5		0.366	0.153/0.579			8.680	.123
Impulse Control [d]	6		0.222	0.040/0.404			3.323	.767

Note: k = number of studies/samples; d_{++} = overall effect size estimate; d_{+} = within category effect size estimate; CI = Confidence Interval; Q_B = between category homogeneity statistic; Q_T = Overall homogeneity statistic; Q_{wi} = within category homogeneity statistic; p=probability level.

[a] The Anxiety category also contains studies originally coded as test anxiety. When left separate those two categories were within .001 effect sizes of each other.

[b] The Other category contains studies on marital help (Phinney, 1977), ethnocentrism (Alsbrook, 1970), prisoner's attitudes (Burt, 1972), happiness (Fordyce, 1977), chronic mental illness (McClaskey, 1970), memory in the elderly (Scogin and Storandt, 1985), and surgery preparation (Young and Humphrey, 1985).

[c] The Self-esteem/concept category included two studies on self-concept (Kingsbury, 1983; Kohutek, 1983) and one on irrational/neurotic thinking (Kassinove, 1980). If the Other and the Self-esteem/concept categories are combined into a broader "Other" category, the resulting $\underline{d}_+ = 0.522$ with $Q_{wi}=19.22$ ($p=.038$).

[d] The Impulse control category contains studies originally coded as alcohol, habit control, and smoking. Because the $\underline{k}$'s on all these Impulse control groups were so small (ranging from 1-3) and the $\underline{d}_+$'s so consistent (ranging from 0 to 0.241) they were condensed into one homogeneous group.

Thus, using problem type as a moderator variable resulted in a significant Q_B and homogeneity within all but three of the resulting subgroups (the Anxiety, Assertion, and Weight Loss categories). An inspection of outliers suggested that the removal of four samples (one from Anxiety, two from Weight Loss, and one from Assertion) and the use of one additional moderator variable within the Assertion category would lead to a parsimonious explanation for the remaining heterogeneity. The rationales, both statistical and non-statistical, for the removal of these outliers are given below.

Mitchell, Hall, and Piatowska's (1975) study was removed from the Anxiety category data because it had the most extreme effect size ($\underline{d}$ = 1.921) and the greatest amount of counselor contact (26 hours of group contact, nearly 2 hours weekly). Jeffrey and Gerber's (1982) and Tobias and McDonald's (1977) samples were removed from the Weight Loss data because of extreme effect sizes ($\underline{d}$= -.296 and $\underline{d}$= 0.000, respectively) and anomalies of design. Jeffrey and Gerber's (1982) negative effect size was actually a condensation of four subsample effect sizes, three of which were positive (mean $\underline{d}$ = +.681 for those three). However, since the largest subsample (self-selected as "inactive correspondent participants") was larger than the other three samples combined and had a negative effect size ($\underline{d}$= -.600), that study's effect size fell to -.296. The Tobias and McDonald (1977) effect size of zero was actually based on a default because of inadequately reported data ($\underline{p}$>.05); their experimental subjects actually did lose an average of

six pounds more than the control group subjects (a finding comparable to other studies with small positive effect sizes). The removal of these three outliers resulted in within group heterogeneity that was explainable by chance for both the Anxiety and Weight Loss categories (see Table 5).

Rakos and Schroeder's (1979) study was removed from the Assertion category data because of its extreme effect size ($\underline{d}$= 2.850). It was also the smallest study in the category ($\underline{N}_e + \underline{N}_c$ =18). After removal of the one outlier in the Assertion category, significant heterogeneity still remained. Further inspection of the data suggested that one other moderator variable, treatment medium, was important in this category. Assertion samples with subjects who received bibliotherapy via an audio-visual medium (typically videotape) had a higher effect size average than those who received bibliotherapy via a book (see Table 5).

Removing this one outlier and subdividing the Assertion category by the moderating variable "treatment medium" resulted in samples that were no longer significantly heterogeneous. These data are summarized in Table 5.

Table 5

Removal of Problem Type Outliers and Subdivision of Assertion Data

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Anxiety (one outlier removed)	14	0.739	0.550/0.930		17.999	.158
Weight Loss (two outliers removed)	7	0.848	0.610/1.080		8.195	.224
Assertion (one outlier removed)	11	0.884	0.673/1.096		26.904	.000
Assertion Subdivided by Bibliotherapy Type [a]				6.215		.013
Audio-visual	7	1.113	0.835/1.390		11.545	.117
Book	5	0.568	0.241/0.894		9.144	.103

Note. [a] The reason there are twelve samples in this treatment medium moderator analysis is that one study (Nesbitt, 1981) contributed two separate samples in this subdivision, one of which received bibliotherapy via a book and one which received audio-visual bibliotherapy. Within the Nesbitt study the audio-visual bibliotherapy had an effect size .218 higher than the book bibliotherapy treatment.

Removing the one Anxiety outlier lowered that category's estimate by one-sixth of an effect size (.167) to .739. Removal of the two Weight Loss outliers dramatically changed that estimate from $\underline{d}_{+}$=.396 to $\underline{d}_{+}$=.848, and the confidence intervals no longer overlap between these two comparisons. Removing the one assertion outlier lowered this estimate from $\underline{d}_{+}$=.946 to $\underline{d}_{+}$=.884.

Thus it appears that problem type is a significant source of effect size variance. The removal of four outliers (5.7% of 70) and the addition of another moderating variable in the Assertion problem type suggested a parsimonious explanation for the remaining variance. However, since it was possible that other models could also explain the heterogeneity of the data, additional analyses were conducted.

Tests of Categorical Models-Dependent Variable Type as Moderator

Effect sizes were also coded for type of dependent variable measure used. This resulted in a total data base of 118 effect sizes from the 70 samples. Obviously, some studies contributed more than one effect size to this analysis because some studies operationalized their dependent variables in more than one way.

The overall effect size estimate ($\underline{d}_{++}$) of this analysis is 0.567, very similar to the original overall effect size estimate of 0.565 in which the 70 studies each contributed only one effect size per study (see Table 6).

The results presented in Table 6 must be interpreted cautiously because they represent nonindependent effect sizes and

because of the large degrees of heterogeneity within five of the seven categories. However, they may suggest that a plausible explanation for the heterogeneity of the bibliotherapy studies was due to dependent variable type (Q_B= 46.407, p=.000).

Table 6

Bibliotherapy Effect Size Estimates: Dependent Variable Types

Sample/Category	k	d_{++}	d_{+}	95% CI	Q_B	Q_T	Q_{wi}	p
Total Sample	118	0.567		0.511/0.623		419.595		.000
Dependent Variable Type					46.407			.000
Nonstandardized Scale	15		1.012	0.829/1.196			47.568	.000
Observed Behavior	20		0.797	0.638/0.956			95.915	.000
Standardized Scale	41		0.564	0.470/0.658			114.335	.000
Academic Achievement	9		0.478	0.317/0.639			38.551	.000
Scale Rating by Other	5		0.436	0.129/0.744			2.743	.740
Physiological	14		0.397	0.242/0.552			52.287	.000
Self Report/Behavior	14		0.329	0.176/0.482			21.789	.083

These results may, however, be confounded by problem type. Fourteen of the fifteen effect sizes (93.3%) in the highest dependent variable type (Unstandardized Self-report) were from the three problem types with the highest effect size estimates (sex dysfunction, assertion, anxiety). Sixteen of the 20 effect sizes (80%) in the second highest dependent variable type (Observed Behavior) came from anxiety and assertion studies. These two highest dependent variable types included no effect size estimates from the lowest scoring problem types of impulse control and weight loss. In contrast, six of the 14 effect sizes (42.9%) in the lowest dependent variable type (Self-reported behavior) came from impulse control studies. Eleven of the 14 effect sizes (78.6%) in the second lowest dependent variable type (physiological) came from either weight loss or impulse control studies.

These results leave a bit of a quandary. It may be that problem types with lower effect sizes (like weight loss and impulse control) scored that way because the types of dependent variables used most often to measure them (self-reported behavior and physiological measures) were prone to lower estimates. It may be that the problem types with higher effect sizes (like sex dysfunction, anxiety,and assertion) scored higher because the types of dependent variables used most often to measure them (unstandardized self-report measures and observed behavior) were prone to higher estimates.

One possible descriptive way of addressing this conundrum

was to graphically compare problem types within each of the dependent variable types to determine if there were consistencies in their relative positions. Separate meta-analyses were made on each dependent variable broken by problem type. The results of these meta-analyses are presented in Table 7.

As can be observed in Table 7, sexual dysfunction, assertion and anxiety tended to have higher effect size estimates regardless of dependent variable type (except for anxiety measured as an observed behavior). Impulse control tended toward a low effect size estimates regardless of its three dependent variable types. It was still difficult to evaluate other problem by dependent variable type confounds because they either rarely had more than two studies within a dependent variable type or generally used only one dependent variable type for that problem (e.g., studying, weight loss). Depression, Career, General, and Other (k's all greater than two) all had moderate outcomes within the standardized self-report dependent variable. More detailed results are presented in Table 19 in Appendix D.

Table 7

Effect Size Comparisons of Problem Types within Dependent Variable Types

	Physio	Observed Behavior	Scale Rate by Other	Standardzd SelfReport	UnStandard SelfReport	Academic Achievmnt	Self-Report Behavior
k->	(14)	(20)	(5)	(41)	(15)	(9)	(14)
d_+							
					SxDsf-1		
1.5							
					Anxty-8		
1.25							
				SxDsf-2			
1.0		Assrt-10		Anxty-11	Career-1		
							SxDsf-4
.75				Assrt-9			
	Assrt-1		Assrt-2				Career-1
		Othr-3		Dprsn-5		Anxty-4	
	Anxty-2				Assrt-5		
.50				Genrl-3			
	Wght-9		Dprsn-1	Wght-1		Study-5	Assrt-2
.25	Impls-2	Anxty-6		Othr-5			Impls-6
				Career-3			
			SxDsf-1				
0.0				Impls-1			Dprsn-1

Note. The numbers in parentheses denote the number of effect sizes (k) in that dependent variable type. The numbers following each problem type are the number of studies of that problem type using that dependent variable type. Within each dependent variable type the problem types are ranked and their approximate d_+ is noted at the far left side. "Anxty" denotes Anxiety studies, "Assrt" denotes Assertion studies, "Dprsn" denotes depression studies, "Genrl" denotes General category, "Impls" denoted Impulse Control studies, "Othr" denotes Other category, and "SxDsf" denotes Sexual Dysfunction studies. More detailed results are in Appendix D.

Tests of Categorical Models: Other Research Methodology Moderators

There may be research methodology variables other than dependent variable type that could moderate these effect sizes. For example, it may be that these effect size estimates are moderated by the type of control group used in the study. Of the 70 studies analyzed, there were 41 that used only No Treatment controls, 19 that used some type of Placebo/Comparison group, and 10 that used both within a study. First, the 41 No Treatment control studies were compared with the 19 Placebo/Comparison studies. Second, a direct within-study comparison was made because this can help to rule out confounds between effect sizes and other study characteristics (Shadish and Sweeney, 1991). This was done by making a direct comparison of the treatment groups effect size difference between No Treatment and Placebo/Comparison groups within the 10 studies that had both. The results of both these analyses are in Table 8.

The results of these two control group analyses are rather congruent with each other. Both show a significant tendency for placebo/comparison group studies to exhibit smaller effect size estimates than No Treatment control studies.

Table 8

Possible Moderators - Control Group Comparisons

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Between Groups Analysis				32.227		.000
No Treatment Controls	41	0.709	0.607/0.810		98.175	.000
Placebo/Comparison	19	0.255	0.136/0.375		48.834	.000
DIFFERENCE		0.454				
Within Groups Analysis				4.131		.042
No Treatment Controls	10	0.885	0.668/1.102		19.889	.030
Placebo/Comparison	10	0.563	0.342/0.785		13.155	.215
DIFFERENCE		0.322				

The between groups analysis suggests a slightly higher effect size difference between the two control group types. This may have been confounded by a non-representative distribution of the problem types within the two types of groups. For example, 19 of the 41 effect sizes (46.3%) in the No Treatment category came from problem types that tended to have higher effect sizes (anxiety, assertion, sexual dysfunction) compared with only 6 of the 19 effect sizes (31.6%) in the Placebo/ Comparison category. In contrast, only 2 of the 41 effect sizes (4.9%) in the No Treatment category came from the problem type with the lowest effect size estimate (impulse control) compared with 3 of the 19 effect sizes (15.8%) in the Placebo/ Comparison category. Because of these confounds, the within-study comparisons may be a more valid estimate.

Meta-analysis has been accused of inappropriately aggregating poor quality studies together with high quality studies. Numerous meta-analysts have attempted to control for this by developing intricate rating systems based upon such things as threats to internal validity and subject assignment strategies. Instead of developing a rating system of questionable reliability and validity, this researcher coded for five objective indicators of potential quality and analyzed them categorically. Those five objective indicators are best phrased in the form of questions: (1) Did the primary investigator(s) have direct contact with the subjects under study? (2) How were the subjects assigned to samples (e.g., random

or otherwise)? (3) From what population were the subjects drawn (e.g., college classroom, traditional referral sources)? (4) Was the study published or unpublished? and (5) Did the study author present the original data as post-test scores or only as gain scores?

Further explanation of these questions is in order. First, if the primary researcher(s) had contact with the subjects (e.g., as the counselor), it could possibly increase the chances of expectancy effects and artificially inflate effect sizes. Second, nonrandom or non-representative assignment of subjects to groups could bias the results. Third, subjects from different population pools could react differently to bibliotherapy.

Fourth, publication type is frequently cited as a potential moderator in meta-analytic research. There is always a possibility that published studies may be prone to having larger effect sizes than unpublished studies because editors are more likely to reject the publication of studies with non-significant results (Light and Pillemer, 1984). Fifth, some studies reported their descriptive and inferential statistics only in terms of gain scores. The current investigator could, therefore, only compute effect sizes from those gain scores. Gain scores are generally considered less reliable than post-test scores (Kerlinger, 1986; Posavac & Carey, 1985).

The results of these analyses are presented in Table 9. In each applicable case the categorical analysis is first presented with missing (i.e., unreported or uncodable) data entered as a separate category, then secondly presented with the missing data deleted.

This missing data were originally left in the analysis in order to determine if they were different from the other categories, but then removed in order not to artificially inflate the probability of Q_B being significant. In the case of the subject assignment question, the three non-random methods of assignment are presented separately and then combined because of small cell sizes.

Table 9

Research Methodology Variables

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Researcher Contact with Subjects?				18.549		.000
Missing	36	0.432	0.338/0.526		125.013	.000
Yes	22	0.717	0.587/0.848		52.987	.000
No	12	0.799	0.605/0.993		27.994	.006
Researcher Contact with Subjects? (Missing data out)				0.472		.472
Yes	22	0.717	0.587/0.848		52.987	.000
No	12	0.799	0.605/0.993		27.994	.006
How were Subjects Assigned?				15.157		.004
Missing	4	0.880	0.430/1.329		15.771	.003
Random	58	0.614	0.535/0.692		165.635	.000
Matched	2	0.269	-0.216/+0.753		2.211	.331
Self-Selected	4	0.184	-0.074/+0.441		25.479	.000
Other Process	2	0.361	0.070/0.651		0.290	.865
How were Subjects Assigned? (Missing data out)				13.233		.004
Random	58	0.614	0.535/0.692		165.635	.000
Matched	2	0.269	-0.216/+0.753		2.211	.331
Self-Selected	4	0.184	-0.074/+0.441		25.479	.000
Other Process	2	0.361	0.070/0.651		0.290	.865
How were Subjects Assigned? (Missing data out/Combined)				12.432		.000
Random	58	0.614	0.535/0.692		165.635	.000
Nonrandom combined	8	0.262	0.083/0.441		28.782	.000

Table 9 continued

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Subject Population Source?				17.210		.002
Mental Health Referral	8	0.843	0.617/1.066		14.104	.079
College Classroom [a]	21	0.632	0.507/0.756		60.679	.000
College Nonclassroom [a]	15	0.620	0.455/0.785		42.259	.000
Noncollege Solicitation	24	0.443	0.324/0.562		90.226	.000
Prison	2	0.101	-0.279/+0.480		0.064	.968
Publication Type				0.449		.503
Published Studies	61	0.557	0.482/0.632		208.757	.000
Unpub. Dissertations	9	0.636	0.417/0.854		15.333	.082
Effect Size Computation				2.744		.098
Gain Score	24	0.650	0.527/0.772		100.383	.000
Post Test Score	46	0.523	0.436/0.610		121.416	.000

Note. [a] The categories of solicitation from College Classroom and College Nonclassroom were coded separately to investigate if possible "perceived coercion" by subjects in classroom settings might make the effect sizes different between the two categories. That is apparently not the case.

Tenatively the results presented in Table 9 suggested several things. First, there appeared to be no effect size differences between the studies where the investigators had contact with the subjects and those studies without such contact. The effect size difference of 0.082 can be attributed to chance (p=.472). Unfortunately, nearly half of the studies did not report sufficiently well for this variable to be coded and the missing data category was significantly different from (i.e., lower than) the reported data. Second, the vast majority of these studies (82.9%) did report using randomization for group assignment and those that did use randomization were significantly higher than those that did not. Third, subject population source may have influenced the overall results because those categories also differed significantly from each other. Subjects garnered from traditional mental health referral sources had the highest effect sizes (albeit with overlapping confidence intervals) and prison subjects had the lowest (but k = 2).

In this analysis publication bias was not apparent. Unpublished studies actually had a slightly higher estimated effect size over published studies, but difference was insignificant (Q_B= .449, p=.503). This must be interpreted cautiously because there were only nine unpublished studies included. In addition, this finding may be confounded by problem type. Four of the nine unpublished studies (44%) were conducted on problem types with higher estimated effect sizes (anxiety and assertion) and none of

the unpublished studies investigated the problem types with typically lower effect sizes (impulse control and weight loss).

Last of all, while gain score computed effect sizes were slightly higher than post-test computed effect sizes, these differences were not significant (p=.098).

In all of the results reported in Table 9, the models are not completely specified because significant heterogeneity (Q_{wi}) remains in some or all categories. Therefore all of these results must be interpreted cautiously.

Tests of Categorical and Continuous Models - Amount of Contact with Therapist.

Two strategies were used to determine if amount of therapist contact moderated the effects of bibliotherapy. First, the studies were broken into three categories: Low contact (< 8 minutes weekly), medium contact (10-29 minutes weekly), and high therapist contact (30 minutes or more weekly). Second, a continuous model analysis (reported below) was run using minutes contact per week as a continuous variable. There were two reasons both a categorical and continuous analysis of this variable were conducted. First, there were twelve studies in which insufficient data was presented to include them in the continuous analysis, but they were able to be reliably placed into one of those three global categories. Second, I believed that having a categorical explanation of contact had a certain simplicity to it that would make understanding this variable more practical. The two analyses should complement each other.

A few of the 70 samples contributed more than one effect size to these analyses because there were groups within studies that varied on their amount of therapist contact. Therefore, this entire portion of the current report must be interpreted cautiously. Three studies provided insufficient data to be coded into one of the three categories.

There were, therefore, a total of 84 effects sizes in the categorical analysis and 72 in the continuous model analysis. Cases in which a study had two samples that differed in amount of therapist contact, but which fell into the same category (e.g., one sample with 45 minutes and another with 90 minutes would both be in high contact category) were combined into one effect size in the categorical analysis, but left separate in the continuous analysis. No study contributed more than two effect sizes to the categorical analysis.

The data in Table 10 suggest that the amount of therapist contact may indeed be a moderating variable; increased therapist contact appeared to be related to increased effect sizes. The homogeneity statistic was significant (Q_B =10.202, p = .006), but the confidence intervals of the groups did overlap. In addition, there was still significant heterogeneity within the categories.

The heterogeneity within the low and high contact categories was investigated by breaking them down by problem type. The effect sizes used in these two separate analyses were independent, that is no sample contributed more than one effect size to the low contact

analysis nor to the high contact analysis. Table 11 summarizes the results of these analyses.

Table 10

Bibliotherapy Effect Size Estimates: Categorical Tests of Amount of Therapist Contact

Sample/Category	k	d_{++}	d_+	95% CI	Q_B	Q_T	Q_{wi}	p
Total Sample	81	0.586		0.516/0.655		246.447		.000
Amount of Therapist Contact					10.202			.006
Low (<8 min)	48		0.506	0.414/0.598			143.428	.000
Medium(10-29)	7		0.805	0.481/1.129			3.919	.789
High (30+ min)	26		0.677	0.565/0.790			91.922	.000

Note: There were three studies that provided insufficient information to be able to be classified into one of the categories above. These three had a d_+=0.153, Q_W=0.231 (p=.972). These three were not included as a discrete category in the above analysis. If the medium and high contact groups had been combined into one category, the following statistics would have resulted: k=33, d_+=0.691, Q_{wi}=96.377 (p=.000), Q_B=6.642 (p=.010), CI = .585/.797.

Table 11

Bibliotherapy Effect Size Estimates and Tests of Problem Type Categories: Low, Medium & High Therapist Contact Samples.

Low therapist contact samples

Sample/Category	k	d_+	95% CI	Q_B	Q_T	Q_{wi}	p
Total Sample	48	0.506	0.414/0.598		143.428		.000
Problem Type				61.199			.000
Sexual Dysfunction	2	2.262	1.413/3.111			0.054	.973
Assertion	10	1.020	0.767/1.274			27.278	.002
Anxiety	13	0.675	0.474/0.877			12.456	.491
Others	6	0.671	0.365/0.976			4.546	.603
Career	3	0.626	0.335/0.918			8.534	.036
Depression	3	0.447	0.030/0.864			4.717	.194
Weight Loss	7	0.164	-0.042/0.369			18.211	.011
Impulse Control	4	0.123	-0.080/0.326			6.433	.169

Medium therapist contact samples.

Sample/Category	k	d_+	95% CI	Q_B	Q_T	Q_w	p
All Med.Contact	7	0.805	0.481/1.129			3.919	.000
Problem Type				1.391			.708
Weight	1	1.212	0.445/1.979			0.000	.999[a]
Anxiety	2	0.768	0.234/1.301			0.013	.993
Sexual dysfunction	2	0.707	-0.032/1.447			0.037	.982
Depression	2	0.653	0.019/1.286			2.479	.290

Table 11 - continued

Sample/Category	k	d_+	95% CI	Q_B	Q_T	Q_{wi}	p
			High therapist contact samples				
Total Sample	26	0.677	0.565/0.790		91.922		.000
Problem Type				49.503			.000
Anxiety	3	1.815	1.425/2.205			2.852	.415
Weight	4	1.039	0.714/1.363			7.424	.115
Depression	2	0.783	0.279/1.287			3.678	.159
Assertion	4	0.692	0.344/1.041			6.830	.145
Other	6	0.561	0.349/0.773			12.085	.060
Impulse Control	3	0.442	0.086/0.799			1.730	.630
Studying	4	0.342	0.123/0.560			7.820	.098

Note. [a] The Q_{wi} statistic requires more than one data point in order to be computed or interpreted.

There were two categories in which high counselor contact samples had significantly higher effect sizes than the low therapist contact samples. They were Anxiety and Weight Loss. There was one category, Assertion Training, in which the increased counselor contact appeared related to lower effect sizes. Other categories either had overlapping confidence intervals or so few effect sizes that it was impossible to identify even tentative trends.

Continuous Variables. To further determine if amount of counselor contact might have a relationship with effect sizes, simple unweighted Pearson correlations were conducted between the effect sizes and relevant continous variables. Scatterplots were drawn to search for potential non-linear relationships. Then Rosenthal and Rubin's focused comparison method (1982) for analyzing whether continuous variables are significant predictors of effect sizes was conducted on variables that were primary research questions or looked to have a potentially predictive value according to the simple Pearson correlations. Weighting studies by sample size is recommended in meta-analysis so larger studies will contribute more to the relationship than smaller studies. The Rosenthal and Rubin method does weight studies by sample size.

The results of the unweighted Pearson correlations are listed in Table 12. The reported probability levels must be held suspect because they do not account for weighting by sample size (Hedges & Olkin,1985).

Table 12

Simple (Unweighted) Pearson Correlations between Effect Sizes and Coded Continuous Variables

	Pearson r	k	p
Length of Treatment (weeks)	0.027	66	.830
Amount Therapist Contact/Weekly	0.034	74[a]	.773
Frequency Therapist Contact/Week	0.016	81[a]	.888

Note: A scatterplot graphically displaying the relationship between Amount of Therapist Contact and effect size can be seen in Appendix F. [a] A few studies contributed more than one effect size to the last two unweighted correlation analyses because some studies had samples within the study that differed on that variable. For example, one study might have two samples, one of which met with a therapist for 30 minutes weekly, the other for zero minutes.

As can be observed in Table 12, the potential predictors did not approach significance levels using unweighted correlations. The Rosenthal and Rubin focused comparison of effect sizes method tests for linear relationships between continuous predictors and effect sizes while weighting studies by sample size (Johnson, 1993). Two-tailed probability levels are reported throughout these continuous model analyses.

Continuous model analyses were conducted on the following variables: Amount of contact with therapist per week, length of treatment (in weeks), and total amount of contact with therapist (in minutes). The total amount of contact with therapist was computed by multiplying the length of treatment in weeks by the amount of contact per week. The results of these analyses in Table 13 are not dissimilar to the unweighted Pearson correlation results reported above. None of the these three continuous variables approached significance levels.

Table 13

Continuous model - Therapist Contact Variables

	z	M	p	k
Amount Therapist Contact	0.397	37.992	.691	74
Length/Treatment	0.622	6.437	.534	82
Total Minutes Contact [a]	1.310	270.617	.190	74

Note. M denotes the mean for that variable (Contact in minutes, length of treatment in weeks) [a] Total amount of contact was computed by multiplying the average amount of therapist contact per week by the length of treatment.

As was stated in the categorical model of Amount of Therapist contact (see above), subjects with different problem types appeared to react differentially to the amount of therapist contact afforded them. The categorical models analysis suggested that the effect sizes in studies on Anxiety, Depression, Impulse Control, and Weight Loss related positively to increased counselor contact, but that effect sizes in studies on Assertion and Sexual Dysfunction might have negative relationships with amount of counselor contact. Other problem types (Career, Studying, Self-Concept/Esteem issues, and Others) had either insufficient numbers or such a restricted range that it was impossible to judge the relationship.

A continuous model analysis of these data should help clarify these relationships. The results of this analysis are presented in Table 14. As was indicated in the categorical analysis, there is evidence that studies dealing with Anxiety and Weight Loss do have a positive relationship with amount of weekly counselor contact. In addition, the effect sizes in Marital, Impulse control, Depression and "Other" categories had positive, albeit nonsignificant, relationships with therapist contact in this continuous model analysis.

The effect sizes in the Studying and Sexual Dysfunction samples had a significant negative relationship with amount of therapist contact. In addition, the effect sizes in the Assertion, Career, and Self-Concept/Esteem issues also had negative, but nonsignificant, relationships with this variable.

The categories in Table 14 are listed with the positive z's first, then the negative z's. Within that breakdown they are ordered from high to low.

The data reported in Table 14 also help highlight another potentially confounding variable in this analysis. The mean amount of time spent with a therapist varies markedly between problem types. Only two of the categories had low mean amounts of contact time (Career and Sexual Dysfunction). Their negative relationships between contact and effect size must be intepreted cautiously, both because of the low numbers in each category (k=3) and their restricted range. If there had been other studies in those categories with therapist contact times of 20 or more minutes, the results of this moderator analysis might have been different.

Table 14

Continuous Model Analysis- Amount of Therapist Contact Moderated by Problem Type

Problem Type	z	M	p	k
Weight	4.112	29.364	.000	11
Anxiety (with Test Anxiety)	3.043	19.100	.002	18
Impulse Control	0.965	31.000	.335	5
Other	0.821	150.000	.411	3
Marital	0.614	45.001	.539	2
Depression	0.232	60.263	.817	8
Studying	-2.563	59.600	.010	5
Sex Dysfunction	-2.113	7.634	.035	3
Career	-1.562	0.517	.118	3
Assertion	-0.485	40.385	.628	13
Self-Concept/Esteem	-0.161	40.001	.872	3

Note. M denotes the mean (in this case minutes per week) for each problem type.

The significant negative relationship between contact time and effect size in the studying category may also be more of a result of a small number (k=5) being influenced by one outlier. Four of the five studies in this category exhibited moderate positive effects (+0.379 to +0.793) with contact times ranging from 0-50 minutes. When viewed in a scatterplot, these four seemed to exhibit only a weak relationship between contact time and effect size. However, Richards and Perri's study (1978) had group contact time of 160 minutes weekly with the subjects and an effect size of -0.253. This negative effect size was an aberration; of the nine studies that used academic achievement as a dependent variable, Richards and Perri (1978) was the only one in which the control group (N=23) improved from pretest to posttest. The control group improvement was slightly more than the treatment group improvement, hence the negative effect size. The control group in this study was a no-treatment control, not a placebo control.

The only two categories where amount of therapist contact did have a significant positive relationship with effect size were Anxiety and Weight Loss. Both of these categories had unrestricted ranges and more studies than the other categories (k's = 18 and 11, respectively), thus making interpretation defendable and plausible.

Tests of Categorical Models - Type of Therapist Contact & Treatment Medium as Moderators

The next two research questions will be addressed together because of their conceptual similarities. Does the type of therapist

contact (e.g., individual, group, mail) moderate the effects of bibliotherapy? Does the medium of the treatment moderate the effects of bibliotherapy, i.e., are self-help books more or less effective than self-help audio-visual presentations?

In both of these cases the Q_B statistic was significant (see Table 15), suggesting that both variables did moderate effect size. It may be, for example, that automated contact is superior to group contact which may be superior to mail contact. But here too we may have problem type confounding these results. Assertion studies disproportionately used automated contact; weight loss and impulse control studies disproportionately used mail contact.

As was stated earlier, Treatment Medium apparently moderated the Assertion studies. It was hypothesized that Treatment Medium might not moderate other studies as much as it did Assertion. It seemed plausible that modeling new behaviors observed on a video or audiotape might have accounted for the increased effect sizes in the Assertion category, but that modeling would not necessarily play as important a role in the other problem types. Therefore the Assertion studies were removed and the data reanalyzed (see Table 15). With the Assertion studies removed the difference between the book and audio-visual mediums was less remarkable and the Q_B statistic nonsignificant.

Table 15

Type of Therapist Contact and Treatment Medium as Moderators

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Type of therapist contact				53.967		.000
Automated	9	1.100	0.877/1.323		20.069	.017
None	3	0.841	0.384/0.130		13.079	.004
Group	33	0.665	0.561/0.769		98.113	.000
Individual	15	0.554	0.348/0.759		18.259	.249
Phone	13	0.522	0.324/0.720		17.601	.173
Other	7	0.479	0.200/0.757		24.800	.001
Mail	8	0.146	-0.014/+0.305		15.185	.056
Treatment Medium				18.154		.000
Book/Manual	54	0.501	0.423/0.578		172.023	.000
Audio-Visual	15	0.931	0.749/1.113		34.206	.003
Treatment Medium (with Assertion Studies Removed)				2.856		.091
Book/Manual	50	0.495	0.416/0.575		162.891	.000
Audio-Visual	8	0.711	0.474/0.948		10.927	.206

Test of Categorical Models - Paradigms as Moderator

The next research question asked whether different theoretical approaches to bibliotherapy might be differentially effective. This was difficult to address because there was considerable overlap in these theoretical strategies within the studies. There were 23 samples judged to use a distinctly behavioral strategy for the treatment, nine that were distinctly cognitive in their approach, and 24 that combined both behavioral and cognitive strategies. Only two studies reported the use of approaches other than behavioral or cognitive. Twelve studies did not report their theoretical strategy.

The results of this analysis are reported in Table 16. The Q_B statistic is significant, suggesting that between groups differences do exist, but considerable heterogeneity remained within the categories.

Cognitive-behavioral and cognitive approaches may, for example, be superior to behavioral approaches. But here too we may have problem type confounding these results. For example, weight loss studies disproportionately (zero of nine) avoided the use of cognitive or cognitive-behavioral strategies; cognitive-behavioral approaches were disproportionately used in anxiety studies (eight of 16) and assertion studies (six of twelve).

Another complication of this analysis is that cognitive, behavioral, and cognitive-behavioral strategies (hereafter referred to as the "Big Three") are all theoretically similar. It would have

been preferable to compare the "Big Three" strategies to, for example, bibliotherapies with a humanistic bent. A few in the "unspecified theory" category may have used theoretical approaches other than the "Big Three", but this was not reported adequately enough to be codable. The two studies within the "varied approaches" category did report group comparisons between one of the "Big Three" and a humanistic approach, but these were insufficient in number to warrant further investigation.

Table 16

Theoretical Approaches as Moderators

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Theoretical Approach to Treatment				28.444		.000
Cognitive-Behavioral	24	0.789	0.662/0.916		92.191	.000
Cognitive only	9	0.709	0.491/0.927		16.856	.051
Varied Approaches	2	0.559	0.127/0.990		0.858	.651
Behavioral	23	0.496	0.376/0.617		58.388	.000
Unspecified	12	0.273	0.118/0.429		27.806	.006

Tests of Categorical Models - Clinical vs. Non-clinical Problems

It was hypothesized that the effects bibliotherapy might have on subjects with clinical problems (e.g., depression, anxiety disorders, sexual dysfunction) would differ from those without clinical problems (e.g., career concerns, assertiveness, self-esteem, smoking, weight loss). Samples that had problems recognized by the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, Third Edition Revised (1987), as clinical disorders were coded as non-clinical if a study indicated that its subjects had only "mild problems" (e.g., problem drinking). Otherwise, client problems defined as mental disorders in the DSM-IIIR were coded as clinical problems; V-codes and problems not defined in the DSM were coded as non-clinical.

The results of this analysis suggested that there were no significant differences in the effects of bibliotherapy between clinical and non-clinical populations (see Table 17).

Tests of Categorical Moderators - Use of Cash Deposits

Another research question addressed whether the presence of a cash deposit, fee, or payment might have a positive relationship with effect size. It was thought that the presence of a cash transaction might motivate subjects to become more seriously involved with their therapy; this might be even more important for bibliotherapeutic interventions since interpersonal motivations tend to be minimized.

Those studies that used a cash transaction did have effect

sizes that averaged slightly higher. However, the Q_B statistic in this analysis was not significant (p=.071, see Table 17) and highly significant heterogeneity remained within the categories. In addition, this analysis may have been confounded by problem type. For example, while only 24.3% of all studies used a cash transactions strategy, a disproportionate number of impulse control studies (66.7%; four of six) used such a strategy.

Table 17

Clinical Problems and Cash Deposits as Potential Moderators

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Clinical Population				0.089		.765
Yes	18	0.545	0.396/0.694		43.249	.001
No or V Code	52	0.571	0.490/0.652		181.205	.000
Cash Deposit/Fee Presence				3.253		.071
Present in study	17	0.698	0.537/0.858		47.034	.000
Not reported present	53	0.533	0.454/0.612		174.256	.000

Failsafe N Estimate

In order to estimate the number of new or unfound studies averaging null results needed to reduce the findings of this meta-analysis to a neglible level (i.e., the "File Drawer Problem" identified by Rosenthal,1979, 1991), a failsafe N was computed using Orwin's (1983) formula. This was done on the main analysis of this report where there were 70 studies, each of which contributed only one effect size. The "neglible level" used in this analysis was +.071 because this is the maximum $\underline{d}_{++}$ that could cause the 95% confidence interval to span zero. In effect this tests how many studies would be required to reduce the $\underline{d}_{++}$ = .565 to a nonsignificant level. The failsafe N was found to be 487 studies, suggesting that 487 studies averaging null results would be necessary to lower the mean effect size of this meta-analysis to .071 or lower.

As it could be argued that some of these 70 effect sizes contained considerable non-bibliotherapeutic treatment (because of the samples with high amounts of counselor contact time), a separate failsafe N was conducted on the 48 effect sizes defined as "low-contact with therapist" (less than 8 minutes weekly). This failsafe N was found to be 216 studies, suggesting that 216 studies (with less than 8 minutes of weekly contact) averaging null results would be necessary to lower the mean effect size of this meta-analysis to a nonsignificant level (in this case $\underline{d}_{++}$=.092).

Direct Comparisons with Therapist-Administered Therapies.

One important research question was "How do bibliotherapies compare to traditional therapist-directed therapies in effectiveness?" It was hypothesized that in direct comparisons therapist-directed treatments might outperform bibliotherapy, especially bibliotherapy studies with low degrees of therapist contact. It was also hypothesized that bibliotherapies with high degrees of therapist contact, because they were essentially therapist-directed treatments with a bibliotherapy adjunct, might outperform therapist-directed treatments with no bibliotherapy.

There were a total of 30 usable studies found that made a direct comparison within the study between bibliotherapy and a purely therapist-directed treatment. Nine of these were studies not used in any of the earlier analyses because they lacked a control group; these nine did, however, allow an effect size to be computed because they made a direct comparison between therapist-directed and bibliotherapies. Five of the samples contributed two effect sizes each to the analysis because each had distinct bibliotherapy groups that differed in their amount of therapist contact time. There were, therefore, a total of 35 effect sizes in this analysis.

Since nine studies included in this analysis did not make a direct statistical comparison between posttest scores of treatment and control groups, the following strategy was used to find the effect sizes for this analysis. Within a study, the therapist-only treatment mean was subtracted from the bibliotherapy group mean

and the difference was divided by a pooled standard deviation. A positive effect size (d) would indicate bibliotherapy's superiority within a study, a negative one indicated a superiority for therapist-directed treatment.

The overall effect size estimate between therapist-only treatments and bibliotherapy was d_{++}= -0.080. The homogeneity test was not significant, Q_T=39.145, p=.507. The 95% confidence interval did overlap zero (C.I. = -0.199/+.040), suggesting that this difference may have simply occurred by chance.

Since the homogeneity statistic was not significant, there was no statistical rationale for searching for moderator variables. This suggested that there were no differences between therapist-directed treatments and bibliotherapies that differed in their amount of therapist contact.

Follow-Up Efficacy

It is possible that bibliotherapy is effective in producing short-term, post-test results, but is there evidence that these results maintain over time?

Twenty-nine of the original 70 studies reported some type of usable follow-up comparison (e.g., the wait-list control group had not been exposed to a therapy during follow-up). However, four of these did not make a follow-up comparison on a variable that they had also used at posttreatment. This resulted in a final database of 25 (35.17% of 70) usable studies in this analysis. The follow-up times ranged from 4 days to 2 years with a Median= 6 weeks,

M=19.05 weeks, s.d.= 28.65. One study simply reported its follow-up results, but did not specify how long the follow-up time period was (Donner & Guerney, 1969). Sixteen of the 24 studies that reported length of time to follow-up (66.7%) had follow-up time periods less than seven weeks. Only five (20.8%) had follow-up time periods greater than 34 weeks.

At post-test, the 25 samples in this analysis had a d_{++}=0.566. This was not significantly different from the other 45 samples that did not report the use of a follow-up measure (d_{++}=0.565, Q_B=.0002, p=0.987). At follow-up the 25 samples in this analysis had a d_{++}=0.342, suggesting a small to moderate erosion of effectiveness beyond posttest. Moderator analysis and homogeneity statistics follow below.

Many studies that reported multiple dependent variables at posttest reported fewer dependent variables at follow-up. Only those posttest variables included as part of the follow-up were used in this comparison process. This helped assure that the effect sizes were directly comparable.

In order to search for moderators, a similar strategy was used as had been in the therapist-to-bibliotherapy analysis. A new effect size directly comparing each sample's posttest result with the follow-up result was computed. Therefore, in this next analysis, samples with positive effect sizes indicated that the follow-up scores were higher than the posttest scores. Negative effect sizes indicated some outcome erosion over time. This strategy was used

in order to minimize the potential effects of extraneous moderator variables on the follow-up effect size estimate.

The overall effect size estimate between posttest and follow-up was -0.224, suggesting there was a small to moderate amount of erosion of posttest effect sizes (confidence intervals of -0.32/-0.13 do not overlap zero). However, the calculation of the homogeneity statistic, Q_T = 164.122, p=.000 indicated significant heterogeneity among the effect size estimates.

Several strategies were used to search for moderators that would explain this heterogeneity. Problem type, amount of contact with therapist (low, medium, or high), length of follow-up time period (low, medium, or high) and dependent variable type were all analyzed as possible moderators; three of these analyses resulted in significant Q_B statistics and the fourth neared significance (p=.058) with significant heterogeneity (Q_{wi}) remaining within categories (see Table 20 in Appendix E for all four analyses). However, all four of these analyses seem to have been influenced by the presence of three outlier effect sizes, all of which came from the same research program (McFall & Lillesand, 1971; McFall & Twentyman, 1973a, McFall & Twentyman, 1973b). With the removal of these outliers, d_{++} was raised from -0.224 to -0.053, and the Q_T was lowered from 164.122 ($p < 0.000$) to 21.019 ($p = .478$). These three outliers were assertion studies with low therapist contact, a short duration between post-test and follow-up (four days to four weeks), and observed behavior as a dependent variable. All three had used

telephone callers making unreasonable requests as a strategy for measuring their outcome. All three had found their subjects to be quite assertive at post-test (d's ranging from 1.736 to 2.554) but rather compliant at follow-up (d's ranging from 0.000 to -0.722).

It is difficult to determine what artifact most influenced these extreme scores. The three outlier studies had a mean unweighted d = -2.118, the three assertion studies that were not outliers had a mean unweighted d =+0.013. Two of the three non-outlier assertion studies (Royce & Arkowitz, 1978; Nesbitt, 1981) used follow-up dependent variables that were self-reported scale ratings rather than observed behavior; the third (Mcfall & Marston, 1970) came from the same research program as the outliers and used the same telephone caller strategy, but did not exhibit the erosion with time that the three outliers did.

With these three outliers removed the resulting 95% confidence interval (around d_{++} = -0.053) was -0.15 to +0.05, suggesting the possibility that the erosion of effect with time may be minimal. This result must be interpreted cautiously, however, because it is based on a relatively small number of studies (k=22) and the removal of 12% of the studies as outliers.

In addition, several problem types were not at all represented in this follow-up analysis. None of the career, depression, general counseling, or sexual dysfunction studies reported any usable follow-up effect sizes. Several of the depression and sexual dysfunction studies reported follow-up results that suggested little

erosion in effect, but had always treated their "wait-list control groups" during the follow-up time period; they were, therefore, not included in this analysis. It should again be noted that most of the follow-up time parameters were short (Median = 6 weeks) and that perhaps the effects do erode over longer time periods.

Investigation of Other Coded Continuous Variables in Main Analysis

There were a variety of other variables coded for in the main analysis. These variables were not among the twelve primary research questions listed at the end of Chapter One. Two of these variables, reading ability of the subjects and readability of the bibliotherapy material, would have been pertinent research questions if enough study authors had reported such data.

The unweighted correlational analysis of 14 of these variables is reported in Table 18. Only two of these variables had significant relationships with effect size; the other twelve did not approach significant levels. The two variables that were significant were "total sample size" (r= -.313) and "subject retention in study" (r=+.268). The variable "total sample size" was inversely related to effect size, suggesting that larger samples tended to have smaller effect sizes. This result would not come as a surprise to those familiar with meta-analytic work. Small samples tend to be biased towards larger effects sizes (Hedges & Olkin, 1985).

The result of the other significant variable, "subject retention in study," was perhaps more counter-intuitive. A continuous model analysis (using the Rosenthal and Rubin method) was conducted on

the retention rate of subjects in study. The result of this analysis was not dissimilar to the unweighted Pearson correlation results reported above ($\underline{z}$ = 3.806, $\underline{M}$ = 88.296, $\underline{p}$=.000, $\underline{k}$ = 72). Both anaylses suggested that there is a significant relationship between retention and effect size, i.e., as retention rates of studies were higher, effect sizes tended higher. If retention rate is regarded as a study quality variable, this would further suggest that poorer quality studies do not necessarily result in higher effect sizes and that perhaps the opposite is true.

Table 18

Simple (Unweighted) Pearson Correlations between Effect Sizes and Other Coded Continuous Variables

	Pearson r	k	p
Publication Date	-0.156	70	.196
Therapist Years/Experience	0.384	6	.452
Subject Age (mean per study)	0.079	39	.632
% of female subjects (per study)	-0.092	62	.479
Education Level (mean per study)	-0.166	45	.277
% Subjects with Some College Ed.	-0.224	41	.159
Reading Ability of Subjects	-0.495	5	.397
Readablity of Biblio Material	0.521	3	.651
% of Subjects Completing Biblio	0.101	28	.608
% of Reading Material Completed	0.118	35	.500
Amount of Cash Deposit	0.238	16	.375
Total Sample Size	-0.313	70	.008
Subject Retention in Study	0.268	72 [a]	.014

[a] A few studies contributed more than one effect size to the last correlation analysis because some studies had samples within the study that differed on that variable.

CHAPTER V

DISCUSSION

According to the results of this meta-analysis, it appears that bibliotherapy does generally have a moderate degree of effectiveness ($\underline{d}_{++}$ =0.565). However, the most important research question for this meta-analysis was not about the overall effectiveness of bibliotherapy. The most important implications of this meta-analysis had to do with the differential effectiveness of bibliotherapy under different conditions and the cautions raised in interpreting these conditions. In general, the following problem types appear to be most amenable to change via bibliotherapy: Assertion (especially if the therapy is administered in audio-visual formats), anxiety (including test anxiety), and sexual dysfunction. Other problem types seemed moderately amenable to change via bibliotherapy: Career, depression, and other types that had insufficient numbers to be categorized. Still other problem types seemed not very amenable to change via bibliotherapy: Weight loss, academic/studying problems (excluding test anxiety), and impulse control problems (i.e., habits, alcohol use, and smoking). Also, these generalizations seemed consistent in situations where there was very little (less than eight minutes weekly) therapist contact with

the subjects. The confidence intervals of the higher three categories (Assertion, anxiety, and sexual dysfunction) did not overlap with the lower three categories (Weight loss, studying, and impulse control).

It may even be possible to describe these group differences in a more general way. It could be argued that people with problems in the three categories with the lowest effect sizes (weight loss, academics, and impulse control) generally would be knowledgeable of at least some commonly used (but perhaps overly simplified) strategies to address their concerns (e.g., eat less, cut out sweets, study in a quiet place, smoke fewer cigarettes). The bibliotherapy strategies may be more sophisticated versions of strategies already suggested to them through, for example, the media or friends. It could also be argued that these three problem types all require greater resolve or "will-power" to break away from deeply ingrained behavior patterns. All three problem types require the client to disengage from behaviors from which they receive rather immediate rewards and engage in behaviors in which gratification will be delayed.

Similarly, it could be argued that the three problem types with the highest effect sizes have a commonality. All three of these are probably less prone to having change strategies known to or discussed by the lay public or media. For example, a person's neighbor may have several suggestions on how to lose weight (and these may not be strikingly different from bibliotherapeutic

suggestions), but not many neighbors would have specific suggestions on how to overcome premature ejaculation or the fear of snakes. The general public may simply know less about how to overcome these problems and, therefore, they may be more able to change when exposed to new, specific recommendations on how to change.

Scogin, Bynum, Stephens, and Calhoon (1990, p. 45) stated that "the majority of (bibliotherapy) studies dealt with rather circumscribed problems that may lend themselves to more education- and information-based interventions." The current analysis would support that contention. Perhaps certain problems, such as sexual dysfunctions, lack of assertiveness, and anxiety, lend themselves to information-based interventions even more than other types.

Limitations in Interpretation within the Meta-Analysis

There are several cautions to these broad generalizations. First, these generalizations are made on relatively small categories, except for anxiety and assertion. Second, if two outlier studies are removed from the weight loss data, then weight loss seemed considerably more amenable to bibliotherapeutic change. Third, it is also difficult to determine from these data alone whether the reason bibliotherapy was less effective with some problem types was due to a bibliotherapy/problem type interaction or whether some problem types are less prone to change regardless of the therapeutic mode. For example, in their meta-analysis of

psychotherapy, Smith and Glass (1977) found anxiety outcome measures to have higher effect sizes than other outcomes and academic measures to have lower ones. Fourth, the overlap of the confidence intervals changed when the outliers were removed. With the outliers removed and homogeneity present in all categories, the only significant differences were that Sexual Dysfunction and Audio-visual Assertion training were higher than Studying and Impulse Control.

The reader is also cautioned that just because a specific problem type seemed amenable to change does not mean every publication of bibliotherapy will produce that change. The current investigation did not attempt to address the relative efficacy of different books within problem types. Ogles, Lambert, and Craig (1991) compared four books about coping with grief and found no differences in effect between them. However, it is highly probable that there are some poorly written bibliotherapies (just as there are some bad therapists) for every one of the problem types investigated. It is also possible that some of the problem types with lower estimated effect sizes may be more amenable to change if a better bibliotherapy becomes available.

In addition, the interpretation of this meta-analysis, like all treatment effect meta-analyses, is limited by the validity of the reports used as its data. The interpretation of problem type is an example of this. Despite the fact that interrater agreement on problem type was perfect, this does not guarantee that the original

study authors adequately operationalized their problem types or used subjects that were truly representative of those problem types. In other words, the researchers of the studies used in this analysis were explicit about what problem types they were attempting to study, but it is impossible to tell if their attempts were adequate. The interpretations about bibliotherapy's differential effectiveness across problem types (and other variables) must, therefore, be cautious.

It was actually rather disconcerting to discover how few of the bibliotherapeutic titles used in these 79 studies had been multiply investigated. I found only two titles investigated three or more times: A new guide to rational living (Ellis & Harper, 1975) and Slim chance in a fat world (Stuart & Davis, 1972). Only six other titles were researched twice, including popular titles like Feeling good: The new mood therapy (Burns, 1980) and Your perfect right (Alberti & Emmons, 1990).

Many of the titles that are commonly prescribed for clients by clinicians (see the section on surveys of clinical use in Chapter 2) were not found to have any empirical support that met the inclusionary criteria of the current meta-analysis. Also, since all of the studies were conducted with populations that responded to some form of solicitation and, therefore, nearly always had some contact with a therapist, we still know very little about the helpfulness of self-help books marketed for bookstore shelves. Even the citations excluded from the current database did not seem to

address measuring the effectiveness of "pure" self-help treatments, i.e, where a lay person purchases a self-help book with no intention of seeing a therapist. At a minimum, the authors and publishers of such books and tapes should include mail-in surveys to determine how the public is responding to their publications.

There may also be an additional complication in interpreting effect size data across problem types and outcome variables. While it has been eloquently argued by many of the meta-analytic developers (e.g., Glass, McGaw, and Smith,1981; Hedges and Olkin, 1985) that the use of effect sizes is an important improvement over combining significance tests and that effect sizes convert diverse data to the same scale, research synthesizers must be careful in interpreting meta-analytic results. For example, a particular effect size estimate may not really mean the same for weight loss as it does sexual dysfunction. A weight loss of 10 pounds may be an important one, but may show up in a meta-analysis as an effect size of only 0.20. because the weight of the subjects in the study is greatly dispersed (e.g., s.d.=50). In contrast, an effect size for premature ejaculation could conceivably be 2.00, but in reality be increasing ejaculatory latency from 30 seconds to an only slightly less discomforting 90 seconds (assuming a s.d. = 30 seconds). The well-publicized study that showed aspirin to be highly effective in preventing heart attacks had an effect size of only .068 (r=.034, sic, Rosenthal, 1991). Meta-analysts and meta-analysis readers may need to be cautious in interpreting effect sizes across problem

types and outcome variables because what may be a very small effect size with one problem may in actuality be a sizable gain in real terms and vice versa.

It may be that type of outcome measure moderated the effect sizes of bibliotherapy, although this model was not sufficient because a significant amount of heterogeneity remained within the resulting categories. The results of this analysis seemed to be confounded by an imbalance of problem types within outcome variable types. However, there was descriptive evidence that assertion, anxiety, and sexual dysfunction categories had higher effect sizes regardless of the type of outcome measure used (except, perhaps, when anxiety is measure as an observed behavior). Impulse control effect sizes tended to be low regardless of the type of outcome measure used. Weight loss and academic problems were measured almost exclusively with only one dependent variable type each. The majority of other problem types were measured with standardized self-report scales.

The effects of bibliotherapy also may be moderated by type of control group. Studies that had placebo control groups tended to have effect sizes that were around one-third of an effect size lower than studies that used only no-treatment control groups. This may suggest that a portion of bibliotherapy's effectiveness is due to expectation effects.

The results of other research methodology variables suggested that they had little moderating impact on effect sizes and that when

they did, studies that were higher in quality tended to have higher effect sizes. Whether the researcher had contact with the subjects, whether the study was published or not, and the method of effect size computation (gain score or post-test score) seemed not to moderate effect sizes. How subjects were assigned to groups (random or otherwise), the subject population source (e.g., college students or traditionally referred clients), and the within-study subject retention rates may be moderators of effect size, but these generally indicated that "better" studies had higher effect sizes. Studies in which the subjects were randomly assigned had higher effects than those studies that did not use random assignment. The "traditionally referred client" category had a higher effect size than any other category. Finally, studies with high subject retention rates tended to have higher effect sizes, suggesting that subject mortality was not generally a threat to internal validity within bibliotherapy studies. Rosen's concern (1987) that bibliotherapy studies might be overly prone to subject mortality did not appear to be validated in this sample of studies. In the 68 (of 79) studies that reported subject retention the mean retention rate was 84.3% (s.d.=14.6%).

It should be noted that significant within-categories heterogeneity remained in these last three potential moderators and, therefore, they must be interpreted cautiously.

The amount of contact time with a therapist did not, in general, relate to effect size. However, there were two problem

types, weight loss and anxiety, in which amount of contact did have significant positive relationships with effect size (p=.000 and p=.002, respectively). Both of these categories had sufficient numbers of study samples to be reasonably confident in the results. Two problem types had significant negative relationships between contact time and effects sizes: Academic problems and sexual dysfunction. However, both of these were based on very low k's and may have been artifactual. All other problem types had either an insufficient number of studies to elicit statistical power or simply may not have had a relationship with therapist contact time.

The medium of the bibliotherapy treatment (i.e., book vs. audio-visual) may moderate its effects. The evidence was rather conclusive that an audio-visual presentation was a more effective way to learn assertiveness skills. Perhaps, in learning to be more assertive, it is important that a subject actually see and hear models acting assertively; reading a transcript of person speaking assertively may not be sufficient. The evidence that audio-visual self-help presentations were more helpful than books in other problem types was much less remarkable.

It was difficult to determine if bibliotherapies using different theoretical strategies differed in their effectiveness. The vast majority of studies in the current meta-analysis used either cognitive, behavioral, or some combination of both as treatment strategies. These strategies are conceptually similar to each other. In addition, the differences between these strategies were

confounded by problem type.

It was surprising that there were no apparent differences between the effectiveness of bibliotherapy and traditional psychotherapy without bibliotherapy. It had originally been hypothesized that therapist-directed therapies with bibliotherapy as an adjunct (arguably categorized here as bibliotherapy with greater than 10 minutes of therapist contact) would outperform therapist-directed therapy without bibliotherapy which would outperform bibliotherapy with little or no therapist contact (arguably defined as less than 8 minutes of contact weekly). However, all three of these categories fell within one-tenth of an effect size of each other with no significant heterogeneity between the study effect sizes. This result was based on a total of 35 studies, nine of which compared bibliotherapy as an adjunct to psychotherapy without bibliotherapy and 26 which compared bibliotherapy with little or no therapist contact with psychotherapy without bibliotherapy (see Appendix H).

This result should not be assumed to suggest that bibliotherapy is as effective as therapist-directed treatments for more severe, clinical concerns since so few of the studies were conducted on clinical populations with clinical concerns. For example, only one of the studies in the therapist-to-bibliotherapy comparison analysis was conducted on a population of clinically depressed subjects (Wollersheim & Wilson, 1991). The bulk of the studies in this particular analysis addressed problem types like

weight loss, career concerns, smoking reduction, or anxiety concerns. None of the anxiety studies addressed serious clinical concerns such as agoraphobia with panic attacks or generalized anxiety disorder, but tried to help with more benign concerns like test anxiety, simple phobias (e.g., spiders, snakes) or social phobias (e.g.. public speaking). This reservation in interpreting the data is especially important for problems like depression where there are potentially serious consequences and ethical considerations for clients who receive inadequate treatment (Craighead, et al., 1984).

Do the effects of bibliotherapy deteriorate after the therapy has ended? The evidence here is mixed. The results of 22 of the 25 studies that provided useful follow-up data can be interpreted as little or no deterioration, at least over the short-time periods (Median = 6 weeks) reported in these studies. However, there were results from three outlier studies on assertion (all from the same research program) that suggested a marked deterioration over relatively short time periods if observed behavior is the outcome variable.

If it is defensible that there is little or no erosion of bibliotherapy's effects at follow-up, perhaps there is a theoretical explanation. It could be that persons who choose (or at least use) self-help strategies like bibliotherapy may be more prone to attributing responsibility for problem solutions to themselves (Brickman, et al., 1982) and, therefore, be more prone to maintaining the effects they feel they have accomplished through their own

strength and effort. This theoretical assertion is, of course, still tenative and in need of empirical support.

Any meta-analysis is dependent upon a reliable and adequately specified coding scheme. There was considerable evidence that the coding scheme used in the current meta-analysis was reliable despite the heterogeneous research designs used in its studies. However, another researcher might code the variables in this heterogeneous dataset differently and discover somewhat different outcomes.

Another related weakness of this meta-analysis is the non-use of the statistic "normative effect size" (NES). The NES allows the researcher to evaluate the clinical significance of the outcomes by comparing the mean of the posttest outcomes of treatment to the mean of a normative group (Durlak and Lipsey, 1991). This NES allows the researcher to determine if an improvement in therapy is a clinically important improvement, i.e., has the treated sample achieved a posttest score that is in or near the realm of "normality?" While the issue of clinical significance of bibliotherapy is an important one, the use of the NES was impractical with this set of studies. None of the studies reported data necessary to compute the NES and because these studies were on such a diverse set of problem types with a diverse set of outcome measures, finding the data necessary would have been extremely time-consuming and probably incomplete.

Similarly, it may have been informative to correct these

effect sizes for reliability of their outcome measures. But here again very few of these studies reported reliability data. The studies that did not report reliability used such a large and diverse set of outcome measures that collecting the data would have been very time-consuming and probably resulted in incomplete data anyway.

Limitations in Interpretation External to the Meta-Analysis

An insufficient number of the studies that met the inclusionary criteria of the current meta-analysis did not evaluate, or at least report evaluating, several potentially important moderating variables. For example, there were only five studies that reported information about the reading ability of the subjects in their studies. It would seem a plausible hypothesis to suggest that reading ability might be positively related to the effects of bibliotherapy. Similar questions include "Does the readability of the bibliotherapy material match the reading ability of the subjects?" and "Is there a positive relationship between education level or cognitive aptitude of subjects and the effects of bibliotherapy?" Only three of the studies in the current meta-analysis reported on the readability of the bibliotherapy text. While it was possible to code education level for 47 of the 79 (59%) studies, only 14 of those 47 truly reported the education level of their subjects. Thirty-three of those 47 were actually the default of 13.1 (freshman in college, 1st month) given to studies known to be conducted on college populations, but which failed to report the actual mean grade level.

Perhaps even more important, no studies reported a within study correlation between education level (nor aptitude or reading ability) and the effects of bibliotherapy.

In addition, these studies inadequately reported bibliotherapy compliance rates and none reported a correlation coefficient between subject compliance and effectiveness. It is recommended that future bibliotherapy evaluations ask subjects to what degree they have read the material and report the reasons why they may not have?

There was also inadequate information available about personality match as a potential moderating variable for bibliotherapy. Only four studies in the current meta-analysis reported some form of these data. In all four of these cases the matching variable was something simplistic like "strength of belief that bibliotherapy would be helpful."

Other studies that were not included in the current meta-analysis (because they lacked a control group) were considerably more illuminating on this variable. For example, Holland Realistic types were found to be more successful and Enterprising types to be less successful at reducing depression with bibliotherapy according to Mahalik and Kivlighan (1988). They also found that clients with high generalized self-efficacy and an internal locus of control benefited most from bibliotherapy. Schallow (1975) also found that subjects who were more successful with self-modification had a significantly higher internal locus of control score than those who

were less successful. Ogles, Lambert, and Craig (1991) found that participants reporting higher expectations of receiving help from bibliotherapy reported greater symptomatic change. Reppucci and Baker (1969) reported that students who were more energetic, outgoing and well-organized, viewing themselves as powerful and competent showed the greatest improvement in self-directed desensitization. More studies comparing bibliotherapy's effectiveness for people who differ in locus of control, Holland types, and/or the "Five-factor model of personality" (e.g., McCrae & Costa (1986) are highly recommended. It would seem plausible, for example, that persons higher in conscientiousness might be more prone to bibliotherapeutic change.

There were an insufficient number of studies that used more indirect or affective approaches rather than direct, step-by-step, problem-solving approaches. These affective approaches are highly touted by professionals in library science, English education, and poetry, but there is an almost complete lack of quantitative empirical evidence for their effectiveness. For example, there is a Journal of Poetry Therapy that has been published by the National Association for Poetry Therapy since 1987. A hand search of its articles led to some interesting reading and a few qualitative evaluations, but no studies that met the inclusionary criteria of the current meta-analysis. While, admittedly, it might be more difficult to operationalize outcome variables in affective bibliotherapy approaches, nonetheless, the methodologies to do so are available

and should be used. Empirical data are needed to either support or refute the many claims that are being made about affective approaches.

Unfortunately, the pace at which empirical evidence for bibliotherapy's effectiveness has been generated has apparently slowed in the past five to eight years. This does not seem to be accompanied by a slowing of availability of bibliotherapies in book stores. This researcher has no adequate hypothesis for the slowing pace of the empirical evidence.

Another limitation of the current study is that it did not address how effective bibliotherapy might be for children. There are studies available for such a meta-analysis and more of those studies used indirect, affective approaches to bibliotherapy.

The findings of the current analysis are somewhat lower than the results of meta-analyses conducted on similar topics by Scogin, Bynum, Stephens, and Calhoon (1990) or by Gould and Clum (1993). Their overall effect sizes were 0.96 and 0.76, respectively. However, when compared to other meta-analyses conducted on psychotherapy and counseling (mean weighted E.S. = .571; Lipsey & Wilson, 1994), the results of this bibliotherapy meta-analysis are rather similar.

Summary

Mahoney (1988) stated that he believed that the success rates for bibliotherapy in current evaluations suggested that it is more effective than no treatment. Schrank and Engels asserted "that

positive recommendations of the value of bibliotherapy exceed available documentation of its usefulness" (1981, p.146). Both assertions could be supported by the results of the current meta-analysis.

In summary, the results of the current meta-analysis suggest that bibliotherapy may be moderately effective for the generally circumscribed problems and populations the empirical studies have addressed to date. It is perhaps less effective for problem types where controlling the need for immediate gratification (e.g., overeating, smoking, procrastinating studying) is important.

However, there are numerous important shortcomings in the available data. There is a paucity of data available to determine if the most commonly prescribed and purchased self-help books actually have positive effects, especially under conditions of no therapist contact. There are very few studies that directly compare different self-help books that address the same problem type. There is an almost complete lack of quantitative empirical evidence on whether indirect/affective bibliotherapy approaches (like poetry therapy) are effective. There is a limited amount of data that suggests bibliotherapy is effective with traditionally referred clients and those with clinical problems, but less evidence available showing its effectiveness with clients suffering from more severe clinical problems. Last of all, it is strongly recommended that more studies be conducted that look at personality type, reading ability, and education level as moderator variables. These data should,

when available, help clinicians and the public maximize the effectiveness of self-help strategies.

With the changes coming in our health care system - the emphases on patient responsibility, lowered costs, and prevention - bibliotherapy could play an important role. The current meta-analysis provides some limited evidence for the effectiveness of bibliotherapy. However, this analysis has also identified some definite holes in our database that must be filled before we can know more specifically for whom and under what conditions bibliotherapy does and does not work.

Perhaps the American Psychological Association should develop a set of guidelines for development of self-help materials analogous to those for psychological test materials (Rosen, 1987). This might not only help assure the quality of such materials, but spur on the other research that is badly needed in this area.

APPENDIX A

CODE SHEETS

APPENDIX A
CODE SHEETS

#_____ 0010- Version 6/23/92a

_____ Author

_____Publication Date

_____Group assignment methods (1Rndm, 2Mtchd, 3SlfSelc, 4Othr,9deflt)

_____Control/Comprsn grp (1NoTreat,2ComprisoPlcbo, 3None,4Othr,9deflt)

_____Presence of strictly therapist-administered (non-biblio) treatment comparison (1yes, 2no)

_____Publicatn Type (1Jrnl; 2Book; 3ERIC; 4ThesDis; 5Presntn; 6other)

X_X_X_XTreatment Type (1SelfAdm; 2MinCo; 3TherapAdm; 4TherapDir)

Amount of Therapist contact (Mean minutes per week - InsData99)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

Length of Assmt/Orientation Contact (Total minutes - InsData99)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

Amount of therapist contact (Frequency of session/weekly) (ID99)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

Type of therapist contact(1IndvFace,2Grp,3Phn, 4Weigh-in, 5Mail, 6none, 7automated, 8other, 9deflt)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

Length of treatment (In weeks; 99 default [months x 4.33])

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

Period between treatment end & evaluation (Weeks; 99 def)

_____ End _____Follow-up1 _____Follow-up2

_____ Training Level of Therapist/Contact Persns (1Paraprof, 2GradStu, 3M.S., 4Ph.D, 5Mixed, 9deflt)

_____ Years of experience of Therapist/Contacts (98varied, 99deflt)
_____ Was researcher(s) also a therapist/contact (1yes, 2no, 9?def/NA)

_____DependentVar#1 _____

_____DependentVar#2 _____

_____DependentVar#3 _____

_____DependentVar#4 _____

_____DependentVar#5 _____

_____DependentVar#6 _____
(1Physio; 2ObsBeh; 3ScaleRatByOthr;4SelfRepVal, 5SelfRepNonVal,6AcadAch,7other,8SelfRepBeh)
X_X_X_X XReported Reliability of Dependent Variable

_____ Problem Type (01 to 14; 99 default) _____

Sample Size: Total _____; 999 default)
_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

_____ Type Reading Material (1 Manual; 2 General Pub)

Length of reading material (998NA, 999default/not reported)
_____ #pgs _____minutes (for audio, video, or computer)

Title of material _____

Author(s) & Date _____

_____Age of Subjects (Reported mean or median; 999 deflt)

_____ Gender Ratio (% Female; 999 default)

X_X_X_XXRace of subjects (1Afro; 2Asia; 3Cauc; 4Hisp; 5Indi; 6Mix; 9default)
_____ Education Level of subjects (In years, 99default)

_____ % with some post-secondary education (999default)

_____ Read. Ability of subjects(Estimated Grade-Equiv; 999 default)
_____ Readability of material (Estimated Grade-Equiv; 999 default)
_____ Personality style Match (1Locus, 2Holland, 3MBTI, 4Other,9dflt)
_____ Treat.Medium (1Paper; 2Audio; 3Video; 4CompInfo; 5CompIntractv)
_____ Instruction Type (1Direct Instruct; 2 Indirect/Affective)
_____ Compliance to reading A (% subjects who completed; 999default)
_____ Compliance to reading B (% of reading completed; 999default)

Mortality (Drop-out) rate: Total _____ (% retained thru treatment end)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

Drop-out rate2: Total _____ (% retained thru follow-up1)

Drop-out rate3: Total _____ (% retained thru follow-up2)

_____ Presence of cash deposit, fee, or payment (1 yes; 2no;9 deflt)

_____ Amount of Deposit (or payment)

_____ Amount of fee

_____ Source of subjects (1SolictCollClass; 2SolctColPop; 3SolictedNonColl, 4TradRefer; 5other; 9 default)

_____ Clinical Population? DSM diagnosis (1Yes, 2 No, 3Vcode, 9deflt)

Severity of problem (norms of dependent variable - pretest)

_____ % overweight
_____ # cigarettes daily
_____ Use of alcohol (#drinks daily)
_____ Depression _____
_____ Phobia _____
_____ Sexual Dysfunction _____
_____ Other _____
_____ Additional _____
_____ Additional _____
_____ Additional _____
_____ Additional _____

_____ Paradigm (1 behavioral; 2 cogntv; 3humanis; 4other; 5cog&beh; 9default/unspecif)

_____ Extra-biblio Homework (1Yes; 2No; 9 default)

Comments_____

_____

_____

Study Author _____

Effect Size computed for_____End ____Follow-up1 ___Follow2

_____ DependentVar# ____ _____
(1Physio; 2ObsBeh; 3ScaleRatByOthr; 4SelfRepVal,
5SelfRepNonVal, 6AcadAch, 7other, 8SelfRepBeh)

g	d	
_____	_____	AllTherpst/Allcontrols
_____	_____	AllBiblio/Allcontrols
_____	_____	AllTherpst/Allbiblio

g	d		g	d		g	d	
____	_____	Thpst1/C1	____	____	Thpst2/C1	____	____	Thps3/C1
____	_____	Thpst1/C2	____	____	Thpst2/C2	____	____	Thps3/C2

g	d		g	d		g	d	
____	_____	Bib1/C1	_____	_____	Bib2/C1	____	_____	Bib3/C1
____	_____	Bib1/C2	_____	_____	Bib2/C2	____	_____	Bib3/C1

g	d		g	d		g	d	
____	_____	Thp1/Bib1	____	____	Thp2/Bib1	____	____	Thp3/Bib1
____	_____	Thp1/Bib2	____	____	Thp2/Bib2	____	____	Thp3/Bib2
____	_____	Thp1/Bib3	____	____	Thp2/Bib3	____	____	Thp3/Bib3

g	d		g	d		g	d	
____	____	Bib1/Bib2	____	____	Bib1/Bib3	____	_____	Bib1/Bib4
____	____	Bib2/Bib3	____	____	Bib2/Bib4			
____	____	Bib3/Bib4						

Check how computed:

____Mean&S.D. ____t-tests ____F-tests _____ chi-square

____proportion ____p-value _____other (name)_____

APPENDIX B

CODEBOOK

APPENDIX B
CODEBOOK

BIBLIOTHERAPY CODING SHEET TRAINING MANUAL

This manual should guide and enable the coder to reliably use the coding sheet for studies. It follows the order of the code sheet. The words in boldface are the explanation of each code sheet section. In some cases 2-3 cells of the code sheet are taken and discussed together.

Version # **(e.g., 6/15/92, 10/24, 10/10) Different versions of this coding sheet were used as the coding process was being fine-tuned. These different versions may be found in the files. The sheet with the latest date (6/23/92a at this time) is the one that should be used.**

#_____ 0010- **Identification #. When all studies have been collected they will each be assigned an ID # to be entered into the computer program.**

_____ Author - Please put the author(s) last name(s).

_____Publication Date (year published)

_____Group assignment methods (1Rndm, 2Mtchd, 3SlfSelc, 4Othr,9deflt) **- The study should report if the groups were randomly assigned, matched on characteristics, self-selected, or perhaps chosen some other way. If this is not reported, code 9. If 4Othr is chosen note in the margin the method chosen. If subjects are stratified by some pertinent characteristic first and then randomly assigned, code 1.**

_____Control/Comprsn grp (1NoTreat,2ComprisoPlcbo,3None, 4Othr,9dflt) **- Code what type of control group, if any, was used. No treatment control groups are coded 1. If the control group received some form of placebo treatment or treatment that the researcher designed to not be effective, code 2. If there was not a control group (common on smoking & weight loss studies) just a pre-post-test measure, code 3. If it is unclear as to how the control group was chosen, code 9. If it is clearly specified, but does not fit in 1-3, code 4 and note in the margin the process used to choose.**

_____Presence of strictly therapist-administered (non-biblio) treatment comparison (1yes, 2no) **- If the researchers compared bibliotherapy treatment to another type of treatment (e.g., group therapy, individual psychotherapy), code 1. Otherwise code 2. If the therapy in question used the book(s) of the bibliotherapy, but was enhanced by individual or group, code 2, but note the enhanced version in the following sections (e.g., amount of therapist contact). If the therapy in question used the same technique(s) described in the book(s), but did not actually provide the subjects with the book, code 1. This coding should be used to determine inclusion/exclusion into later cells of "ThrpstAdm" and "Biblio."**

_____Publicatn Type (1Jrnl; 2Book; 3UnpubOthr;4ThesDis; 5Prsntn;6othr) - **Was this study published in a journal, code 1; from a book, code 2. If the study was unpublished but did not come from a thesis or dissertation (e.g., an ERIC microfiche), code 3. If it came from an unpublished thesis or dissertation, code 4; if from a presentation at a professional meeting, code 5. Code 6 if other.**

X_X_ X XTreatment Type (1SelfAdm; 2MinCo; 3TherapAdm; 4TherapDir)
- This coding is not used in the later versions of the code sheet. Do NOT code. It was simply impossible to operationally define each of these types.

Amount of Therapist contact (Mean minutes per week - InsData99)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

- If reported, code the number of minutes per week that the therapist was in contact with the client. If the contact was designed to compare bibliotherapy with a traditional individual or group approach, code the number of minutes of the traditional approach in the ThrpstAdm cells. If the therapy included having the clients read a book or otherwise participate in some form of self-administered therapy, code in the Biblio cells.

If the therapist-administered approach used the same techniques used in a bibliotherapy approach it was compared with (or another approach) but did not actually provide the clients with reading material, code it in the ThrpstAdm cells. There may be more than one type of bibliotherapy or therapist administered; if so code each separately (if possible). Do NOT include time used for assessment or orientation to the study; that should be coded in the next section. If clients meet with a therapist more than once per week, be sure to multiply the length of the sessions by the appropriate number. EXAMPLE: If clients meet with 125the therapist for two 90 minute sessions for 3 weeks, then for one 60 minute for 7 weeks, the mean # of minutes coded would be:

$$\frac{(2 \times 90 \times 3)+(60 \times 7)}{10 \text{ weeks}} = 101 \text{ min.}$$

If no exact number of minutes is given, but a SMALL range is, code using the median between the range estimates. SMALL range here should be defined as a top estimate equal to or less than 1.5 of the bottom estimate, e.g., 30-45 minutes would be a small range with a median of 37.5. A range wider than that should be coded Insufficient data 99 (e.g., 15-60 minutes). This definition of SMALL range should be used on other similar time estimates throughout the coding (e.g., 6-9 months, 20-30 years old).

If the number of weeks of the treatment is not cited, but the total number of minutes of treatment is (e.g. 2 hours = 120 minutes), put the number of minutes in the cell.

Length of Assmt/Orientation Contact (Total minutes -Intake98;InsData99)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

- Often the studies used in this meta-analysis report that there was an initial "assessment" or "orientation" contact with bibliotherapy participants. These contacts vary in length and purpose. Some are simply to present the materials to the participants and to pretest in some manner. Others seem be longer and to have a psychoeducative quality about them. There seem to be such a number of them that to ignore this variable or to include it as part of the previous code ("Amount of therapist contact in minutes") would be problematic. Therefore, if the author of a study describes an initial contact as "orientation" or something similar (not as therapy), code the length of that contact in minutes. Typically the same number would be entered in all the cells of ThrpstAdm, Biblio, and Comparison. If the study states that there was an initial "intake interview" of unspecified length, code 98. Be sure not to include this time in previous cells of "amount of THERAPIST contact in minutes."

Amount of therapist contact (Frequency of session/weekly)(ID99)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

- Typically the number to be entered into the cells here is "1" session per week. However, some studies had therapist contact that was more or less frequent. If a therapist had contact (e.g., by phone, in person) with the therapist biweekly, put 0.5 in the cell (.33 if triweekly, etc). If the therapist met more frequently, indicate so with an appropriate number above 1. If the subjects had contact with the treatment center, but not a therapist (e.g., they read their bibliotherapy materials at the treatment center library or worked on an automated device during specific appointments), code amount of therapist contact as 0 (see compliance cells for further instructions). If the therapist contact varied as the therapy progressed, compute the number as was done in the "Amount of therapist contact in minutes." For example, if clients meet with the therapist twice weekly for 3 weeks and then once weekly for 7 weeks, the frequency would be coded:

$$\frac{(2\times3)+(1\times7)}{10 \text{ weeks}} = 1.3 \text{ sessions weekly}$$

Type of therapist contact(1IndvFace,2Grp,3Phn, 4Weigh-in, 5Mail, 6none, 7automated, 8other, 9deflt)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

- Studies vary according to their type of therapist contact. It is perhaps impossible to code 6none because there almost must be some type of communication between the participants and the researcher (unless the subjects were participating without their knowledge and the dependent variables were measured in a purely unobtrusive way). The other five codes are relatively frequent. This should be interpreted as the most typical type of contact between the therapist and participant and should not include the contact made in any initial "orientation" session. If subjects had primary contact with an automated or semi-automated device (computer or interactive audio system) code 7. If the therapist contact is something other than those described, code 8. If the report does not adequately define the type, code 9.

Length of treatment (In weeks; 99 default [months x 4.33])
_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

127 **- Typically authors report the length of the treatment in weeks. In some cases the length of client self-paced bibliotherapies may be given in the mean or median length to completion. If the time is reported in months multiply by 4.33; if given in days divide by 7. If a study reports the number of sessions involved in the treatment, but does not explain the timing of those sessions sufficiently for the earlier cells (i.e., frequency of sessions weekly) to be coded, simply code the number of sessions in these "length of treatment" cells.**

Period between treatment end & evaluation (Weeks; 99 def; months x4.33)
_____ End _____Follow-up1 _____Follow-up2
- Typically the number that will go in the "END" cell will be 0; however, occasionally a researcher may gather the data on the dependent variable a few weeks later. The studies may report on the length between the "END" data collection and a follow-up evaluation (an evaluation to determine if the treatment effects were consistent or improving with time).

_____ Training Level of Therapist/Contact Persns (1Paraprof, 2GradStu, 3M.S., 4Ph.D, 5Mixed, 9deflt)
- This is often not reported (9deflt). Code "2GradStu" even if the therapists appear to have a master's degree finished but are currently in training beyond their master's (e.g., interns, practicum students). Code 5mixed if there is a variety of therapists reported to have worked with the participants (e.g., "2 Ph.D.'s, 3 interns, and 2 practicum students"). Code 1Paraprof if the therapist has a bachelor's degree or less and is not an active graduate student in training.

_____ Years of experience of Therapist/Contacts (98varied, 99deflt)
- Usually an author may report the mean or median years of experience of the therapists involved. If only a range of experience is reported (e.g., 2-12 years experience) code 98.

_____ Was researcher(s) also a therapist/contact (1yes, 2no, 9?def/NA)

- Usually the author will mention whether or not she was involved in the therapy. If even one researcher was also a therapist, code 1.

_____DependentVar#1 _____

_____DependentVar#2 _____

_____DependentVar#3 _____

_____DependentVar#4 _____

_____DependentVar#5 _____

_____DependentVar#6 _____

(1Physio; 2ObsBeh; 3ScaleRatByOther; 4SelfRepVal, 5SelfRepNonVal, 6AcadAch, 7other, 8SelfRepBeh)

- Studies vary widely on the type of dependent variables measured for outcomes. In addition, some studies have a variety of variables. Code them in the order that they are presented in the study and label them in the space provided so that reliability checks can be accurate. It is also important to label them so a check can be run on the frequency of use of different measures (e.g., the Beck Depression Inventory, various anxiety scales). If a dependent variable is measured physiologically (e.g., weight loss, biofeedback, amount of nicotine in the blood stream) code 1. If the d.v. is behavior observed BY THE RESEARCHERS (e.g., social skills or approaching a phobic object like a snake) code 2.

If the d.v.is a rating scale of an inferred construct scored by a person other than the subject (e.g., improvement in self-esteem, anxiety level, attitude change) code 3. This rating will be relatively infrequent on studies done on adults, but probably more frequent on studies done with children. If the d.v. is measured with a self-reported measure that has validity/reliability data available (e.g., MMPI, Fear Survey Schedule, Beck Depression Inventory, Primary Communication Inventory) code 4. If the d.v. is measured with a self-reported measure that appears not to be validated by previous research (e.g., "on a scale of 1-10 rate how anxious you feel now" or an instrument that was designed specifically for this research project and not yet validated), code 5. If the d.v. is measured by academic achievement (e.g., GPA, ability or achievement test scores) code 6. If the d.v. is behavior observed and reported by the subject alone (e.g., smoking cessation, hours spent studying, number of drinks consumed) code 8 - NOTE the order change. If the d.v. seems to be measured in a way that does not fit with the other 7 codes, code 7other and briefly describe in the margin how the d.v. was measured.

In a few cases researchers measured variables for which they did not expect there to be a treatment effect (e.g., a generalization effect from snake phobia treatment for spider phobia d.v.'s). These are usually cases analagous to convergent/divergent validity checks. Code ONLY those variables for which there WOULD be an expected treatment effect.

X X X XReported Reliability of Dependent Variable
129 - This was an earlier code that is no longer used. If a study reports the reliablity of a measure (e.g., internal consistency of .94) note that in the margin with the appropriate d.v.

_____ Problem Type (01 to ; 99 default) _____

CODE: 01 Alcohol
02 Anxiety
03 Assertiveness
04 Career Indecision
05 Depression
06 General Counseling
07 Habit control (not smoking)
08 Heart disease
09 Marital/couple dissatisfaction
10 Sexual dysfunction
11 Smoking
12 Studying problems
13 Test Anxiety
14 Weight
15 Other
16 Heterogeneous problems

The grouping of these problem types is a judgement call and the validity of inclusion of a study within a category is a judgement call. This coding is perhaps the most subjective of the entire process. This will be a coding where reliablity checks will be especially important. These groupings come from this researcher's survey of nearly 100 of the studies to date. Anxiety disorders includes specific phobic disorders, speech anxiety, generalized anxiety disorder, agoraphobia, and general stress management, but not test anxiety. Later on the form there will be a code about whether or not the population met specific DSM diagnostic criteria. General counseling includes self-esteem and self-concept issues, but not anxiety disorders, depression, or other problems coded separately.

Habit control will typically be nail-biting, but other behavioral habits (other than smoking and compulsive overeating) can be included. This category should NOT include habits described in the DSM classification as disorders (e.g., kleptomania, obsessive compulsive disorder, pathological gambling, bulimia nervosa). The marital/couple dissatisfaction and sexual dysfunction categories seem to be mutually exclusive in this literature. Studying problems includes all forms of academic performance other than cognitive ability and test anxiety.

The coder should state in the space provided what the label is. If 15other is coded, a brief explanation should be included.

Sample Size: Total _____; 999 default)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

- This should be the sample size at the end of treatment, not beginning. The retention rate for the study, including follow-ups, is coded later on the form. If the sample sizes are broken down by cell, code each cell as well as the total; if only a total is given, place in that box and note how many groups the total was divided into.

_____ Type Reading Material (1 Manual; 2 General Pub)

- General Publication ("2") here would include materials that were initially designed for the popular self-help press, not as a part of a research program. Examples would include Codependent no more (Beattie, 1987), Feeling

good(Burns, 1980) or Your perfect right(Alberti & Emmons, 1990). "Literature" (e.g., poetry or fiction) used as bibliotherapy should also be coded 2. Other materials that were originally designed as part of a research program should be coded 1. Automated programs should be coded 1. Originally this code was to differentiate between published and unpublished materials. However, this was problematic because a number of the bibliotherapy materials were unpublished in the earlier research and published in later studies. The published versions of research programs generally have a step-by-step structure to them; general publications (popular press) generally do not.

Length of reading material (998NA, 999default/not reported)
_____ #pgs _____ minutes (for audio, video, or computer)
- Generally the article will include the length of the material either in pages or minutes. Computer programs may not have a set length in minutes, but may be reported in the mean number of minutes per client. Code 998 in the cell that is not appropriate, 999 if the length is not reported.

Title of material _____

Author(s) & Date _____
- Simply print the requested information in the spaces provided.

_____ Age of Subjects (Reported mean or median; 999 default)
130 **- Self-explanatory. Report to one decimal place.**

_____ Gender Ratio (% Female; 999 default)
- If reported, take the the number of women and divide by the total number of subjects (report the percentage the closest whole number). For example if a study had 44 women and 28 men, the number entered would be 61 (%). A few studies may be 100% or 0%.

_____Race of subjects (1Afro; 2Asia; 3Cauc; 4Hisp; 5NatAm; 6Mix; 9defaut)
- It appears that no studies reported this variable, but if they do use the codes above. If there is a breakdown (e.g., 20% Asian, 30% Native American), code 6 and write the breakdown in the margin.

_____ Education Level of subjects (In years, 99default)

_____ % with some post-secondary education (999deflt)
- Various studies report education level differently, if at all. If a study reports the average grade level (e.g., 14.5 for college students), code that in the "Education Level" cell. If the study reports how many were at each level of education, add the percentages of those with some post-secondary education (degree unfinished, B.A., graduate degree) together and code that in the latter cell. Both cells may be coded if the information is available. If the study used college students as subjects, but did not clarify their grade level, put 13+ in the former cell and 100 in the latter.

_____ Read. Ability of subjects(Estimated Grade-Equiv; 999 default)

_____ Readability of material (Estimated Grade-Equiv; 999 default)

_____ Personality style Match (1Locus, 2Holland, 3MBTI, 4Other,9dflt)

- These 3 codes above would seem to be critical in our understanding of the differential effects of bibliotherapy. The empirical questions "Do subjects of various personality styles or reading abilities have differential outcomes with bibliotherapy?" seem to be largely ignored in the literature. If the information is provided, report it in the appropriate cell. For personality style match 1 is for locus of control, 2 for a Holland (RIASEC) code, 3 for a Myers-Briggs code, and 4 for anything else. Defaults will be the most common for these cells. If any of these are coded a brief description in the margin is warranted (e.g., this study was done on Holland code "Investigatives" only).

_____ Treat.Medium (1Paper; 2Audio; 3Video; 4CompInfo; 5CompIntractv)

- Most of the bibliotherapy research has been done in the treatment medium of books (i.e., 1Paper). However, if the study reports bibliotherapy materials in audio, videotape, informational computer, or interactive computer, code accordingly. If a study uses two or more treatment mediums, code either 6 for paper plus a technological medium or 7 for a combination of technological mediums.

_____ Instruction Type (1Direct Instruct; 2Indirect/Affective)

- Indirect/Affective bibliotherapy would include poetry therapy, the use of fiction, the use of non-fiction stories analogous to the problems a person might be having, etc. With these the emphasis is on the metaphorical, affective nature of the reading. Direct instruction would generally suggest practical strategies and information in a step-by-step format to help with problems (e.g., cognitive and behavioral strategies for weight loss or smoking cessation). Automated programs should be coded 1 unless they are primarily poetry, fiction, etc.

_____ Compliance to reading A (% subjects who completed; 999default)

_____ Compliance to reading B (% of reading completed; 999default)

- Compliance is reported in two different ways, if at all, in this literature. While these would appear to be critical factors in the effectiveness of bibliotherapy, relatively few researchers report about treatment compliance. In the former cell, report what percentage of subjects reportedly completed the bibliotherapy. In the latter cell, report what the mean or median percentage of reading was completed by each subject. A few studies may report enough data to code both cells. Generally the data will have to be computed into percentages, especially for the former cell.

If subjects read their bibliotherapy or worked on automated devices at appointed times at the treatment/research center or as part of a class, code 100 in each of the cells unless the study states otherwise.

Mortality (Drop-out) rate: Total _____ (% retained thru treatment end)

_____ThrpstAdm1 _____TherpstAdm2 _____TherpstAdm3

_____Biblio1 _____Biblio2 _____Biblio3

_____Compariso1 _____Compariso2

Drop-out rate2: Total _____ (% retained thru follow-up1)

Drop-out rate3: Total _____ (% retained thru follow-up2)

- Researchers may report mortality rates directly (i.e., in percentages) or may provide the data with which to compute the percentages (e.g., 86 people were screened and actually started the therapy, but only 64 participated in the post-treatment assessment = 74%). Some studies will report that a larger number of persons inquired about participation, but only a select number of those met the criterion for the study and actually began the therapy or participated in the control group. Use only the number who met the criterion and actually participated as the base number from which to compute the mortality rates. Use the same base number to compute the percentages for the follow-up assessments as well, e.g., 86 began the study, 64 completed (74%), 58 participated in the 3 month follow-up (67%), and 55 in the 6 month follow-up (64%). Report in whole percentages.

If the study reports mortality rates within each cell, report that information in the cells provided. This will allow the meta-analytic research to ascertain if there are different mortality rates for bibliotherapy versus traditional therapy. If mortality rates are not broken down per cell, report them in the "Total" cells. Report in the "Total" cells for follow-up mortality rates.

_____ Presence of cash deposit, fee, or payment (1 yes; 2 no;9 deflt)

- A number of studies required participants to put down a cash deposit or charged a fee (usually to increase compliance and reduce mortality). If the study stated that there was such a monetary transaction, code 1. If the study STATED that there was NOT such a transaction, code 2. If no mention is made of a monetary transaction, code 9. While it is probably safe to assume that if no mention is made of a monetary transaction, there probably was not one, 9 should be coded unless the study specifically stated that there was not one.

_____ Amount of Deposit (or payment) (999 dflt)

_____ Amount of fee (999 dflt)

- If money is returned to the participant after COMPLETION of therapy requirements, then it is a deposit. If it is not, then it should be coded a fee. For example, if participants were required to deposit $100 at the beginning of the study, but were returned only $75 after completing all the therapy (but not necessarily meeting all their therapeutic goals), then 75 should be coded in the deposit cell and 25 in the fee cell. If money was not returned because a participant did not fully comply with the therapy or attend the assessment, that should be defined as a deposit.

A few studies provided payment to subjects at the end of the study. Code payment in the deposit cell.

_____ Source of subjects (1SolicitedCollClass; 2SolicitedCollPop; 3SolictedNonColl, 4TradRefer; 5other, 9 default)

- It appears that subjects for these studies came from 3-4 basic sources. If the subjects were part of a college course that was solicited to participate in a research study (possibly as part of a course "requirement" for general psychology), code 1. If the subjects came from a college population, but were solicited via campus advertisements, code 2. If the subjects were solicited via advertisement in the general community, code 3. If the subjects were already a part of a clinical population (e.g., clients of a community mental health center) who participated without being solicited through advertisement, code 4. Traditional referral (TradRefer) means that they may already have been receiving treatment and were simply asked if they wished to participate in another form of treatment under study. If a study does appear to have used subjects without informed consent (e.g., subjects who were not aware they were involved in a research project), note that in the margin. If the study found clients in another method besides the 4 previously mentioned, but described its method of solicitation, code 5 (other). If a study does not report how it gathered its subjects, code 9. If a study had subjects from more than one of the sources, code all pertinent numbers (e.g., subjects from both a college class and a college population at large).

_____ Clinical Population? DSM diagnosis (1Yes, 2 No, 3Vcode, 9deflt)

- It appears that most of the RESEARCH literature about bibliotherapy has been conducted on non-clinical populations. If the study used subjects with DSM diagnostic nomenclature, code 1. The disorders most common to watch for are anxiety disorders, psychosomatic problems, depression, sexual dysfunctions, and even "neurotic disorders" (for earlier studies). DSM criteria need not be used, but the term "DIAGNOSIS" (or a derivative) should be used to describe the severity of the subjects' problems. Subjects need not come from a "Traditional Referral" to meet this criterion.

If the study states that the subjects were chosen because of their "mild" forms of a diagnosable disorder (e.g., mild, non-clinical depression), code 2. This "code 2" will also include subjects in smoking, weight loss, assertiveness studies, etc. It appears that most studies on reducing alcohol consumption with "problem drinkers" choose subjects with only "mild" problems, not enough to meet diagnostic criteria for dependence. Subjects with academic problems, marital problems, occupational problems, uncomplicated bereavement, or other "V codes" should be coded 3 unless the study states that their problems were "mild" (then code 2). Code prison populations as 3 (V - Adult Antisocial behavior) unless there is evidence presented otherwise. There are not specific criteria for these V codes.

Severity of problem (norms of dependent variable)

_____ % overweight
_____ # cigarettes daily
_____ Use of alcohol (#drinks daily)
_____ Depression _____
_____ Phobia _____
_____ Sexual Dysfunction _____
_____ Other _____
_____ Additional _____
_____ Additional _____
_____ Additional _____
_____ Additional _____

- Regardless of whether or not the problem being addressed is a clinical one, it is of interest to know how bibliotherapy may work differentially with persons with different severities of problems. Some studies report the norms for their subjects. Three of the most common and directly assessable are percentage overweight, number of cigarettes consumed daily (# packs x 20), and the number of alcoholic beverages consumed daily. For other problems, the severity may be assessed by how far from the norm on some instrument the subjects scored. If possible note how far subjects were from the norm in terms of standard deviations. The "Other" line is included for those problems not listed above. The "additional" lines are included for problems that were measured in more than one way (e.g., pre-treatment depression measured with 3-4 different instruments).

_____ Paradigm (1 behavioral; 2 cogntv; 3humanis; 4other; 5cog-beh; 9default/unspecif)

- Most studies use a bibliotherapy approach fashioned from one of the major psychological paradigms. Most of the studies found for this meta-analysis use either cognitive and/or behavioral techniques. Sometimes the researchers explicitly state that they are using techniques from a certain paradigm. However, if it is obvious to the coder that techniques come from a specific paradigm without an explicit statment from the original researcher, code using that judgement. For example, a study that emphasizes "reinforcement of appropriate behaviors" would be coded 1. A study mentioning techniques of covert desensitization, problem-solving skills, and/or thought-stopping would be coded 2. If a study mixes elements of both 1 & 2, code 5. If it is unclear what paradigms the techniques came from, code 9.

_____ Extra-biblio Homework (1Yes; 2No; 9 default)

- Some studies require subjects to participate in homework activities other than the bibliotherapy itself (e.g., journaling, relaxation training, self-monitoring). If this study requires this type of homework, code 1. If no mention is made of such homework, code 2.

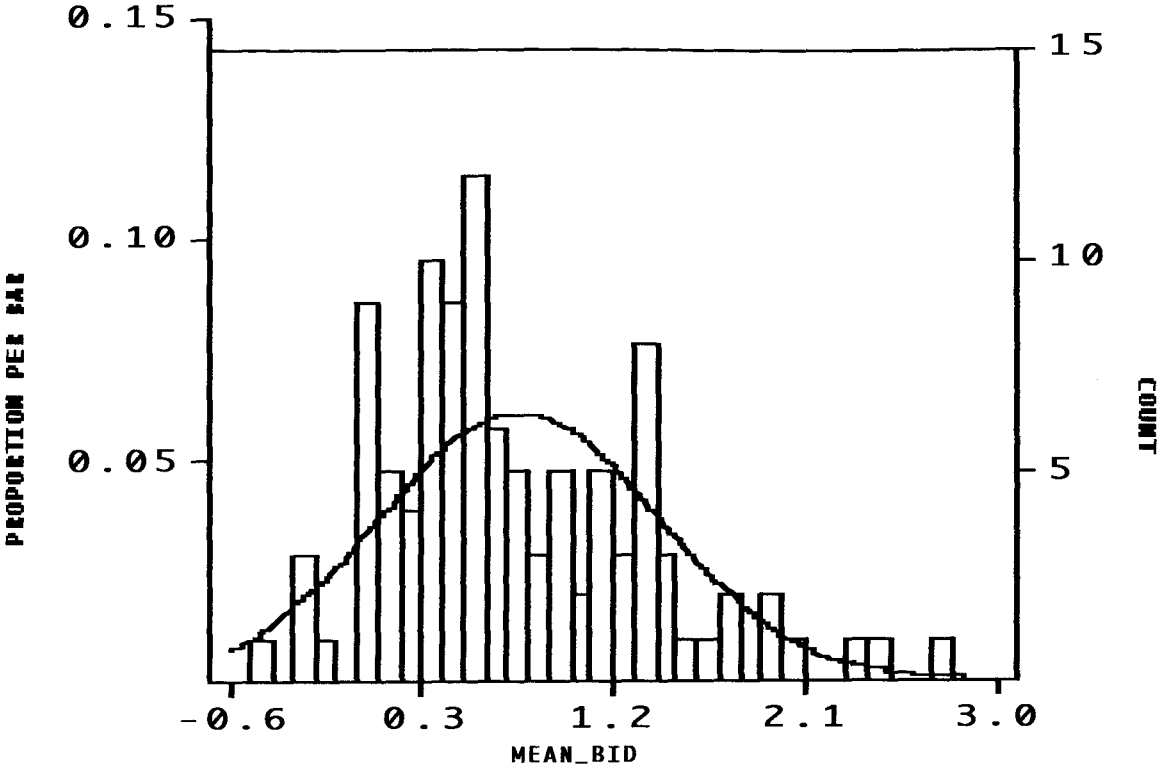
Comments_____

_____

APPENDIX C

GRAPHIC DISTRIBUTION OF 70 EFFECT SIZES

APPENDIX C - GRAPHIC DISTRIBUTION OF 70 EFFECT SIZES COMPARED TO THE NORMAL CURVE



APPENDIX D

Table 19

APPENDIX D

Table 19

Moderator Analyses of Dependent Variable Types by Problem Types

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Physiological D.V.	14	0.397	0.242/0.552	1.418		.701
Assertion	1	0.670	-0.003/1.342		0.000	1.000[a]
Anxiety	2	0.597	0.054/1.140		6.044	.049
Weight	9	0.371	0.191/0.551		43.312	.000
Impulse Control	2	0.301	-0.137/0.739		1.514	.469
Observed Behavior D.V.	20	0.797	0.638/0.956	27.888		.000
Marital	1	2.007	1.364/2.649		0.000	1.000[a]
Assertion	10	1.016	0.780/1.252		58.060	.000
Other	3	0.612	0.259/0.966		1.998	.573
Anxiety	6	0.318	0.020/0.616		7.969	.240
Scale Rating by Other	5	0.436	0.129/0.744	2.243		.524
Assertion	2	0.683	0.229/1.138		0.500	.779
Depression	1	0.401	-0.589/1.391		0.000	1.000[a]
Marital	1	0.203	-0.322/0.728		0.000	1.000[a]
Sex Dysfunction	1	0.150	-0.804/1.104		0.000	1.000[a]
Self-Report/Validated	41	0.564	0.470/0.658	40.175		.000
Sex Dysfunction	2	1.017	0.254/1.781		0.251	.882
Anxiety	11	0.997	0.791/1.203		19.549	.052
Assertion	9	0.780	0.557/1.002		21.869	.009
Depression	5	0.649	0.317/0.980		7.562	.182
General Counseling	3	0.515	0.133/0.896		6.322	.097
Career	3	0.279	0.010/0.547		1.549	.671
Other	8	0.261	0.101/0.422		17.058	.030

APPENDIX D

Table 19- continued

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
SelfReport/Nonvalidated	15	1.012	0.829/1.196	14.111		.003
Sex Dysfunction	1	1.669	0.563/2.774		0.000	1.000[a]
Anxiety	8	1.352	1.060/1.644		31.758	.000
Career	1	1.030	0.630/1.429		0.000	1.000[a]
Assertion	5	0.587	0.284/0.890		1.699	.889
Academic Achievement	9	0.478	0.317/0.639	2.500		.114
Anxiety (Test Anx Only)	4	0.629	0.382/0.877		27.371	.000
Studying	5	0.366	0.154/0.579		8.680	.123
Self-Reported Behavior	14	0.329	0.176/0.482	7.007		.136
Sex Dysfunction	4	0.959	0.384/1.534		11.042	.026
Career	1	0.662	-0.197/1.520		0.000	1.000[a]
Assertion	2	0.389	-0.078/0.855		0.621	.732
Impulse Control	3	0.291	-0.074/0.509		1.517	.678
Depression	1	0.000	-0.528/0.528		0.000	1.000[a]

Note: Cases where there were sub-categories with only 1 effect size may bias the Q_B statistic by artificially increasing the degrees of freedom. In each of the above situations the data was recomputed by combining the k=1 categories into the "Other" or a Small Categories" category. In no case did it change the probability level of the Q_B statistic by more than .02.

[a] The Q_{wi} statistic requires more than one data point in order to be computed or interpreted.

APPENDIX E

Table 20

Appendix E

Table 20

Bibliotherapy Follow-up Effect Size Differences with Moderator Investigations

Sample/Category	k	d_+	95% CI	Q_B	Q_T	Q_{wi}	p
All Studies	25	-0.244	-0.319/-0.130		164.122		.000
Problem Type				65.021			.000
Assertion	6	-1.038	-1.266/-0.810			85.716	.000
Anxiety&Test Anx.	7	-0.243	-0.438/-0.048			10.150	.180
Weight	3	-0.155	-0.533/+0.224			1.057	.788
Studying	4	+0.031	-0.157/+0.218			0.833	.934
Impulse control	4	+0.059	-0.140/+0.257			1.345	.854
Other	1	+0.007	-0.437/+0.451			0.000	.999[a]
Dependent Variable Type				97.089			.000
Observed Behavior	4	-1.560	-1.846/-1.274			49.673	.000
Academic Achmt	5	-0.116	-0.276/+0.044			9.573	.088
Physiological	4	-0.103	-0.412/+0.206			1.272	.866
SelfReport/Valid	5	-0.095	-0.320/+0.131			5.406	.368
SelfReport/Behavior	2	-0.086	-0.495/+0.324			0.547	.761
SelfReport/Nonvalid	2	-0.058	-0.593/+0.477			0.457	.796
Combo of DV's	3	+0.109	-0.112/+0.329			0.081	.994
Amount of Contact With Therapist				5.706			.058
Lo Contact(<8 minutes)	15	-0.340	-0.479/-0.202			147.356	.000
Medium Contact(10-29)	1	-0.274	-0.783/+0.234			0.000	.999[a]
Hi Contact(30+minutes)	8	-0.105	-0.241/+0.031			10.953	.204

Appendix E

Table 20 - continued

Sample/Category	k	d_+	95% CI	Q_B	Q_T	Q_{wi}	p
Duration until follow-up measurement				25.961			.000
<2.1 weeks	4	-0.997	-1.308/-0.687			61.105	.000
4-6 weeks	12	-0.168	-0.296/-0.040			63.062	.000
10+ weeks	8	-0.128	-0.288/+0.033			12.228	.141

Note: All of these moderator analyses were confounded by the presence of the three outlier assertion studies described in Chapter 4. With those 3 removed, overall heterogeneity was explainable by chance (Q_{wi} = 21.019, p = .458). Negative numbers in these tables indicate erosion in effect size between posttest and follow-up, i.e., higher posttest results than follow-up. In each subtable the category with the highest negative number contains the three outliers except for the last analysis (Duration of follow-up measure) where one of the outliers is in the second group.

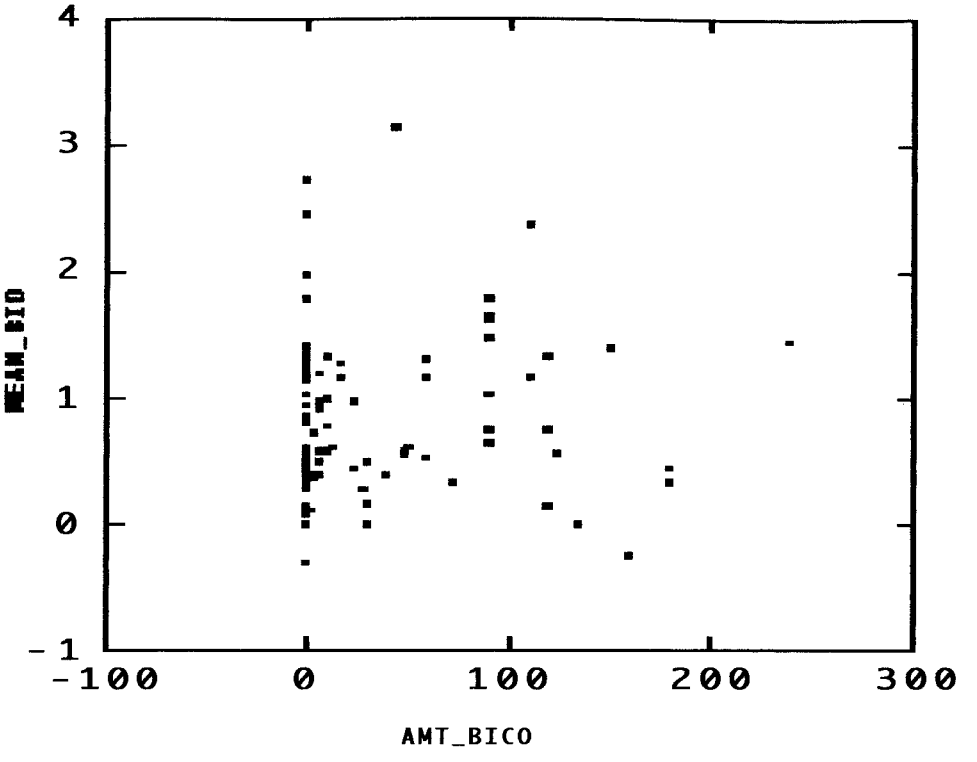
[a] The Q_{wi} statistic requires more than one data point in order to be computed or interpreted.

APPENDIX F

Scatterplot of Amount of Therapist Contact and Effect Size

Appendix F -

Scatterplot of Amount of Therapist Contact and Effect Size



APPENDIX G

Table 21

APPENDIX G

Table 21

Comparison of Therapist-directed vs Bibliotherapy with Moderator Analysis

Sample/Category	k	d_+	95% CI	Q_B	Q_{wi}	p
Overall	35	-0.080	-0.199/+0.040		39.145	.507
Amount of therapist contact				0.571		.752
Lo contact(<8 min wk)	26	-0.073	-0.211/+0.065		33.942	.272
Med contact	3	-0.271	-0.787/+0.245		1.022	.796
Hi contact (30+ min)	6	-0.054	-0.325/+0.217		3.610	.729

Note. The negative effect sizes reported in this table indicate a very small (non-significant) advantage of therapist-directed strategies over bibliotherapy strategies.

APPENDIX H

STUDIES INCLUDED IN THE META-ANALYSES

APPENDIX H

STUDIES INCLUDED IN THE META-ANALYSES

Allen, G.J. (1973). Treatment of test anxiety by group administered and self-administered relaxation and study counseling. Behavior Therapy, 4, 349-360.

Alsbrook, E.Y. (1971). Changes in the ethnocentrism of a select group of college students as a function of bibliotherapy (Doctoral dissertation, University of Illinois, Urbanna-Champaign, 1970). Dissertation Abstracts International, 31, 4536A. (University Microfilms No. 71-05, 024)

Baker, B.L., Cohen, D.C. & Saunders, J.T. (1973). Self-directed desensitization for acrophobia. Behavior Research and Therapy, 11, 79-89.

Balch, P., & Ross, A.W. (1974). A behaviorally oriented didactic group treatment of obesity: An exploratory study. Journal of Behavior Therapy and Experimental Psychiatry, 5, 239-243.

Beneke, W.M., & Harris, M.B. (1972). Teaching self-control of study behavior. Behaviour Research and Therapy, 10, 35-41.

Black, M.J. (1981). The empirical evaluation of a self-administered conversational training program. Dissertation Abstracts International, 42, 1596B. (University Microfilms No. 81-20, 247)

Brown, R.A., & Lewinsohn, P.M. (1984). A psychoeducational approach to the treatment of depression: Comparison of group, individual, and minimal contact procedures. Journal of Consulting

& Clinical Psychology, 52, 774-781.

Brusman, M.M. (1985). A comparison of EMG biofeedback and guided imagery upon anxiety reduction in high-risk college students. Unpublished doctoral dissertation, Northern Arizona University.

Burt, L.N. (1973). Bibliotherapy: Effect of group reading and discussion on attitudes of adult inmates in two correctional institutions (Doctoral dissertation, University of Wisconsin-Madison, 1972). Dissertation Abstracts International, 33, 5211A. (University Microfilms No. 72-31, 519)

Conner, C.N. (1981). The effectiveness of bibliotherapy on teaching initiator dating skills to females. Dissertation Abstracts International, 42, 3818B. (University Microfilms No. 82-03, 109)

Cotler, S.B. (1970). Sex differences and generalization of anxiety reduction with automated desensitization and minimal therapist interaction. Behaviour Research and Therapy, 8, 273-285.

Cuevas, J.L. (1984). Cognitive treatment of chronic tension headache. Dissertation Abstracts International, 46, 53955B. (University Microfilms No. 85-10, 012)

Danaher, B.G. (1977). Rapid smoking and self-control in the modification of smoking behavior. Journal of Consulting and Clinical Psychology, 45, 1068-1075

Dodge, L.T., Glasgow, R.E., & O'Neill, H.K. (1982). Bibliotherapy in the treatment of female orgasmic dysfunction. Journal of Consulting & Clinical Psychology, 50, 442-443.

Donner, L., & Guerny, B.G. (1969). Automated group desensitization

for test anxiety. Behaviour Research and Therapy, 7, 1-13.

Fordyce, M.W. (1977). Development of a program to increase personal happiness. Journal of Counseling Psychology, 24, 511-520.

Glasgow, R.E., Schafer, L., & O'Neill, H.K. (1981). Self-help books and amount of therapist contact in smoking cessation programs. Journal of Consulting and Clinical Psychology, 49, 659-667.

Glasgow, R.E. (1978). Effects of a self-control manual, rapid smoking, and amount of therapist contact on smoking reduction. Journal of Consulting and Clinical Psychology, 46, 1439-1447.

Glasgow, R.E., Swaney, K., & Schafer, L. (1981). Self-help manuals for the control of nervous habits: A comparative investigation. Behavior Therapy, 12, 177-184.

Graff, R.W., Danish, S., & Austin, B. (1972). Reactions to three kinds of vocational-educational counseling. Journal of Counseling Psychology, 19, 224-228.

Hagen, R.L. (1974). Group therapy vs. bibliotherapy in weight reduction. Behavior Therapy, 5, 222-234.

Hanson, R.W., Borden, B.L., Hall, S.M., & Hall, R.G. (1976). Use of programmed instruction in teaching self-management skills to overweight adults. Behavior Therapy, 7, 366-373.

Heather, N., Robertson, I., MacPherson, B., Allsop, S., & Fulton, A. (1987). Effectiveness of a controlled drinking self-help manual: One-year follow-up results. British Journal of Clinical Psychology, 26, 279-287.

Heather, N., Whitton, B., & Robertson, I. (1986). Evaluation of self-help manual for media-recruited problem drinkers: Six month follow-up results. British Journal of Clinical Psychology, 25, 19-34.

Hoffman, M.A., Spokane, A.R., & Magoon, T.M. (1981). Effects of feedback mode on counseling outcomes using the Strong-Campbell Interest Inventory. Journal of Counseling Psychology, 28, 119-125.

James, J.E., & Hampton, B.M. (1982). The relative efficacy of directive and nondirective treatment in behavioral weight control. Behavior Therapy, 13, 463-475.

Jarmon, D.G. (1972). Differential effectiveness of Rational-Emotive therapy, bibliotherapy, and attention-placebo in the treatment of speech anxiety. Dissertation Abstracts International, 33/09-B, 4510. (University Microfilms No. AAD73-06217)

Jeffery, R.W., Danaher, B.G., Killen, J., Farquhar, J.W., & Kinnier, R. (1982). Self-administered programs for health behavior change: Smoking cessation and weight reduction by mail. Addictive Behaviors, 7, 57-63.

Jeffery, R.W., & Gerber, W.M. (1982). Group and correspondence treatments for weight reduction used in the multiple risk factor intervention trial. Behavior Therapy, 13, 24-30.

Kassinove, H., Miller, N., & Kalin, M. (1980). Effects of pretreatment with rational-emotive bibliotherapy and rational-emotive

audiotherapy on clients waiting at a community mental health center. Psychological Reports, 46, 851-857.

Keller, J.F., Croake, J.W., & Brooking, J.Y. (1975). Effects of a program in rational thinking on anxieties in older persons. Journal of Counseling, 22, 54-57.

Kingsbury, S.R. (1983). An analysis of the effectiveness of group psychotherapy and bibliotherapy in changing the self-concept of a selected group of adult married females. (Doctoral dissertation, Southwestern Baptist Theological Seminary, 1983). Dissertation Abstracts International, 45, 339B.

Kirsch, I., & Henry, D. (1979). Self-desensitization and meditation in the reduction of public speaking anxiety. Journal of Consulting and Clinical Psychology, 47, 536-541.

Knauss, M.R., Jeffrey, D.B., Knauss, C.S., & Harowski, K. (1983). Therapeutic contact and individual differences in a comprehensive weight loss program. The Behavior Therapist, 6, 124-128.

Kohutek, K.J. (1983). Bibliotherapy within a correctional setting. Journal of Clinical Psychology, 39, 920-924.

Krivatsky, S.E., & Magoon, T.M. (1976). Differential effects of three vocational counseling treatments. Journal of Counseling Psychology, 23, 112-118.

Lang, P.J., Melamed, B.G., & Hart, J.A. (1970). A psycho-physiological analysis of fear modification using an automated desensitization procedure. Journal of Abnormal Psychology, 76, 220-234.

Lindstrom, L.L., Balck, P., & Reese, S. (1976). In person vs. telephone treatment for obesity. Journal of Behavior Therapy and Experimental Psychiatry, 7, 367-369.

Lowe, J.C., & Mikulas, W.L. (1975). Use of written material in learning self-control of premature ejaculation. Psychological Reports, 37, 295-298.

Mahoney, M.J. (1974). Self-reward and self-monitoring techniques for weight control. Behavior Therapy, 5, 48-57.

Marshall, W.L., Presse, L., & Andrews, W.R. (1976). A self-administered program for public speaking anxiety. Behaviour Research and Therapy, 14, 33-40.

McClaskey, H. (1970). Bibliotherapy with emotionally disturbed patients: An experimental study (Doctoral dissertation, University of Washington, 1970). Dissertation Abstracts International, 31, 5205A. (University Microfilms No. 71-08, 523)

McFall, R.M., & Lillesand, D.B. (1971). Behavior rehearsal with modeling and coaching in assertion training. Journal of Abnormal Psychology, 77, 313-323.

McFall, R.M., & Marston, A.R. (1970). An experimental investigation of behavior rehearsal in assertive training. Journal of Abnormal Psychology, 76, 295-303.

McFall, R.M., & Twentyman, C.T. (1973). Four experiments on the relative contributions of rehearsal, modeling and coaching to assertion training. Journal of Abnormal Psychology, 81, 199-218.

Meyers, A.W., Cuvillier, C., Stalgaitis, S., & Cooke, C.J. (1980). An

evaluation of self-help treatment programs for weight loss. The Behavior Therapist, 3, 25-26.

Miller, W.R. (1978). Behavioral treatment of problem drinkers: A comparative outcome study of three controlled drinking therapies. Journal of Consulting and Clinical Psychology, 46, 74-86.

Mitchell, K.R., Hall, R.F., & Piatrowska, O.E. (1975). A group program for the treatment of failing college students. Behavior Therapy, 6, 324-336.

Monti, P.M., Fink, E., Norman, W., Curran, J., Hayes, S., & Daldwell, A. (1979). Effects of social skills training groups and social skills bibliotherapy with psychiatric patients. Journal of Consulting and Clinical Psychology, 47, 189-191.

Moss, M.K., & Arend, R.A. (1977). Self-Directed contact desensitization. Journal of Consulting and Clinical Psychology, 45, 730-738.

Nesbitt, E.B. (1978). Comparison of two measures of assertiveness and the modification of non-assertive behaviors (Doctoral dissertation, University of Tennessee, 1977). Dissertation Abstracts International, 38, 7236A. (University Microfilms No. 78-07, 713)

Nesbitt, E.B. (1981). Use of assertive training in teaching the expression of positively assertive behavior. Psychological Reports, 49, 155-161.

Ober, D.C. (1968). The modification of smoking behavior. Journal of Consulting and Clinical Psychology, 32, 543-549.

Phillips, R.E., Johnson, G.D., & Geyer, A. (1972). Self-administered systematic desensitization. Behaviour Research and Therapy, 10, 93-96.

Phinney, A.L. (1977). The effects of assertion training and bibliotherapy with married women. Dissertation Abstracts International, 38/06-B, 2879. (University Microfilms No. AAD77-27470)

Pitkow, A.B. (1980). Effects of Rational-Emotive therapy, Rational-Emotive bibliotherapy, and psychodynamically oriented psychotherapy on adjustment in adult clinical outpatients. Dissertation Abstracts International, 42/01-B, 31387. (University Microfilms No. AAD81-13914)

Pyle, K.R., & Stripling, R.O. (1976). The counselor, the computer and career development. Vocational Guidance Quarterly, 25, 71-75.

Rakos, R.F., & Schroeder, H.C. (1978). The development and empirical evaluation of a self-administered assertiveness training program. Journal of Consulting and Clinical Psychology, 47, 991-993.

Register, A.C., Beckham, J.C., May, J.G., & Gustafson, D.J. (1991). Stress inoculation bibliotherapy in the treatment of test anxiety. Journal of Counseling Psychology, 38, 115-119.

Richards, C.S. (1975). Behavior modification of studying through study skills advisement and self-control procedures. Journal of Counseling Psychology, 22, 431-436.

Richards, C.S., McReynolds, W.T., Holt, S., & Sexton, T. (1976). Effects of information feedback and self-administered

consequences on self-monitoring study behavior. Journal of Counseling Psychology, 23, 316-321.

Richards, C.S., & Perri, M.G. (1978). Do self-control treatments last? An evaluation of behavioral problem-solving and faded counselor contact as treatment maintenance strategies. Journal of Counseling Psychology, 25, 376-383.

Richards, C.S., Perri, M.G., & Gortney, C. (1976). Increasing the maintenance of self-control treatments through faded counselor contact and high information feedback. Journal of Counseling Psychology, 23, 405-406.

Rosen, G.M., Glasgow, R.E., & Barrera, M., Jr. (1976). A controlled study to assess the clinical efficacy of totally self-administered systematic desensitization. Journal of Consulting and Clinical Psychology, 44, 208-217.

Rosen, G.M., Glasgow, R.E., & Barrera, M., Jr. (1977). A two-year follow-up on systematic desensitization with data pertaining to the external validity of laboratory fear assessment. Journal of Consulting and Clinical Psychology, 45, 1188-1189.

Royce, W.S., & Arkowitz, H. (1978). Multimodal evaluation of practice interactions as treatment for social isolation. Journal of Consulting and Clinical Psychology, 46, 239-245.

Schelver, S.R., & Gutsch, K.U. (1983). The effects of self-administered cognitive therapy on social-evaluative anxiety. Journal of Clinical Psychology, 39, 658-666.

Schmidt, M.M., & Miller, W.R. (1983). Amount of therapist contact

and outcome in a multi-dimensional depression treatment program. Acta Psychiatrica Scandinavica, 67, 319-322.

Schmidt, M.M. (1980). Amount of therapist contact and outcome in a multimodal depression treatment program. Dissertation Abstracts International, 42/08-B, 3441. (University Microfilms No. AAD82-01952)

Scogin, F., Hamblin, D., & Beutler, L. (1987). Bibliotherapy for depressed older adults: A self-help alternative. Gerontologist, 27, 383-387.

Scogin, F., Jamison, C., & Gochneaur, K. (1989). Comparative efficacy of cognitive and behavioral bibliotherapy for mildly and moderately depressed older adults. Journal of Consulting and Clinical Psychology, 57, 403-407.

Scogin, F., Storandt, M., & Lott, L. (1985). Memory-skills training, memory complaints, and depression in older adults. Journal of Gerontology, 40, 562-568.

Skutle, A., & Berg, G. (1987). Training in controlled drinking for early-stage problem drinkers. British Journal of Addiction, 82, 493-501.

St. Lawrence, J.S., Hughes, E.F., Goff, A.F., & Palmer, M.B. (1983). Assessment of role-play generalization across qualitatively different situations. Journal of Behavioral Assessment, 5, 289-307.

Trudel, G., & Laurin, F. (1988). The effects of bibliotherapy on orgasmic dysfunction and couple interactions: An experimental

study. Sexual and Marital Therapy, 3, 223-228.

Trudel, G., & Proulx, S. (1987). Treatment of premature ejaculation by bibliotherapy: An experimental study. Sexual and Marital Therapy, 2, 163-167.

Wollersheim, J.P., & Wilson, G.L. (1991). Group treatment of unipolar depression: A comparison of coping, supportive, bibliotherapy, and delayed treatment groups. Professional Psychology: Research and Practice, 22, 496-502.

Young, L., & Humphrey, M. (1985). Cognitive methods of preparing women for hysterectomy: Does a booklet help? British Journal of Clinical Psychology, 24, 303-304.

Zeiss, R.A. (1978). Self-directed treatment for premature ejaculation. Journal of Consulting and Clinical Psychology, 46, 1234-1241.

Note: The are actually 80 citations in this appendix. These 80 citations provided the data for the 79 distinct samples used in this meta-analysis.

This discrepancy can be explained as follows: Eight of the citations were actually on only four distinct samples. This occurred because two of the citations (Heather, Robertson, MacPhereson, Allsop, & Fulton, 1987; and Rosen, Glasgow, & Barrera, 1977) were actually follow-up reports from earlier research projects (Heather, Whitton, & Robertson, 1986; and Rosen, Glasgow, & Barrera, 1976). Two additional citations were dissertations (Nesbitt, 1978; Schmidt, 1980) that were found in published form as well (Nesbitt,

1981; Schmidt & Miller, 1983). In all of these situations, both forms of the studies provided unique data for coding, but each distinct sample contributed only one effect size. For example, the published form of Nesbitt (1981) provided enough data to compute the effect sizes, but Nesbitt's dissertation (1978) provided other codable data not in the published version.

Two of the citations actually provided information on five distinct samples. McFall and Twentyman (1973) was actually a report of four distinct experiments, three of which provided usable data for the meta-analysis. Jeffery, Danaher, Killen, Farquhar, and Kinner (1982) conducted two distinct experiments, one on weight reduction and the other on smoking cessation on two distinct samples. Therefore, of the 80 usable citations, 70 had only one sample each, two had a total of five samples, and eight contributed a total of only 4 unique samples. This resulted in a grand total of 79 samples usable in the meta-analysis. Nine of these were used only in the therapist-to-bibliotherapy comparison and 70 were used in the main analysis.

REFERENCES

Alberti, R.E., & Emmons, M.L. (1990). Your perfect right: A guide to assertive living. San Luis Obispo, CA: Impact Publishers.

Burns, D.D. (1980). Feeling good: The new mood therapy. New York: Signet.

Alston, E.F. (1962). Bibliotherapy and psychotherapy. Library Trends, 11, 159-176.

American Psychiatric Association. (1987). Diagnostic and statistical manual of mental disorders (3rd ed. - rev.). Washington, D.C.: Author.

American Psychological Association. (1989). First annual golden fleece awards for do-it-yourself therapies. Presentation at the 97th annual convention of the American Psychological Association, New Orleans, LA.

American Psychological Association. (1978). Report of the task force on self-help therapies. Washington, D.C.: Author.

Anstett, R.E., & Poole, S.R. (1983). Bibliotherapy: An adjunct to care of patients with problems of living. Journal of Family Practice, 17, 845-853.

Atwater, J.M., & Smith, D. (1982). Christian therapists' utilization of bibliotherapeutic resources. Journal of Psychology and Theology, 10, 230-235.

Beatty, W.K. (1964). A historical review of bibliotherapy. Library

Trends, 11, 106-117.

Beck, A.T. (1989). Love is never enough: How couples can overcome misunderstandings, resolve conflicts, and solve relationship problems through cognitive therapy. New York: Harper & Row.

Beck, A.T. (1979). Cognitive therapy for depression. New York: Guilford Press.

Bolles, R.N. (1988). What color is your parachute? Berkeley, CA: Ten Speed Press.

Brickman, P., Rabinowitz, V.C., Karuza, J. Jr., Coates, D., Cohn, E., & Kidder, L. (1982). Models of helping and coping. American Psychologist, 37, 368-384.

Brown, J. (1987). A review of meta-analyses conducted on psychotherapy outcome research. Clinical Psychology Review, 7, 1-23.

Brown, E.F. (1975). Bibliotherapy and its widening applications. Meutchen, NJ: Scarecrow Press, Inc.

Burns, M.T. (1992). Developing biblio/poetry therapy services in a hospice setting: Potential for use with families, volunteers, and the bereaved. Unpublished master's thesis, Concordia University, River Forest, IL.

Burt, L.N. (1973). Bibliotherapy: Effect of group reading and discussion on attitudes of adult inmates in two correctional institutions (Doctoral dissertation, University of Wisconsin-Madison, 1972). Dissertation Abstracts International, 33, 5211A. (University Microfilms No. 72-31, 519)

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

Cooper, H.M. (1989). Integrating research: A guide for literature reviews. (2nd ed.) Newbury Park, CA: Sage Publications.

Craighead, L.W., McNamara, K., & Horan, J.J. (1984). Perspectives on self-help and bibliotherapy: You are what you read. In S.D. Brown & R.W. Lent (Eds.) Handbook of counseling psychology. (pp. 878-929). New York: John Wiley & sons.

Dobson, K.S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. Journal of Consulting and Clinical Psychology, 57, 414-419.

Durlak, J.A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. Psychological Bulletin, 110, 204-214.

Durlak, J.A. & Lipsey, M.W. (1991). A practitioner's guide to meta-analysis. American Journal of Community Psychology, 19, 291-332.

Eagly, A.H., & Johnson, B.T. (1990). Gender and leadership style: A meta-analysis. Psychological Bulletin, 108, 233-256.

Ellis, A., & Harper, R. (1975). A new guide to rational living. Hollywood: Wilshire Books.

Ellis, A., & Knaus, W.J. (1977). Overcoming procrastination. New York: Signet Publications.

Forest, J.J. (1988). Self-help books. American Psychologist, 43, 599.

Fromm, E. (1956). The art of loving. New York: Harper & Row.

Glasgow, R.E. & Rosen, G.M. (1978). Behavioral bibliotherapy: A review of self-help behavior therapy manuals. Psychological Bulletin, 85, 1-23.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills, CA: SAGE Publications, Inc.

Gould, R.A., & Clum, G.A. (1993). A meta-analysis of self-help treatment approaches. Clinical Psychology Review, 13, 169-186.

Groveman, A.M., Richards, C.S., & Caple, R.B. (1975). Literature review, treatment manuals, and bibliography for study skills counseling and behavioral self-control approaches to improving study behavior. JSAS Catalog of Selected Documents in Psychology, 5, 342-343 (Ms. No. 1128).

Hartmann, D.P., Ed. (1982). Using observers to study behavior. San Francisco: Jossey-Bass.

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

Holtje, H.F. (1988). Comment on Rosen. American Psychologist, 43, 600.

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

Johnson, B.T. (1993). DSTAT 1.10: Software for the meta-analytic review of research literatures. [Computer program]. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Johnson, B.T., Mullen, B., & Salas, E. (1993). Hedges & Olkin vs. Rosenthal & Rubin vs. Hunter, Schmidt, & Jackson: A comparison of three meta-analytic approaches. Manuscript submitted for publication.

Kerlinger, F.N. (1986). Foundations of behavioral research. New York: CBS College Publishing.

Kinney, M.M. (1946). Bibliotherapy and the librarian. Special Libraries, 37, 175-180.

Lazarus, A.A. (1988). Right aim, wrong target. American Psychologist, 43, 600.

Libman, E., Fichten, C.S., & Brender, W. (1985). The role of therapeutic format in the treatment of sexual dysfunction: A review. Clinical Psychology Review, 5, 103-117.

Lipsey, M.W. & Wilson, D.B. (1994). The efficacy of psychological, educational and behavioral treatment. American Psychologist, 48, 1181-1209.

Mahalik, J.R. & Kivlighan, D.M. (1988). Self-help treatment for depression: Who succeeds? Journal of Counseling Psychology, 35, 237-242.

Mahoney, M.J. (1988). Beyond the self-help polemics. American Psychologist, 43, 598-599.

Masters, W.H., & Johnson, V.E. (1970). The pleasure bond. Toronto: Bantam Books.

Matt, G.E. (1989). Decision rules for selecting effect sizes in meta-analysis: A review and reanalysis of psychotherapy outcome

studies. Psychological Bulletin, 105, 106-115.

McCrae, R.R. & Costa, P.T. (1986). Personality, coping, and coping effectiveness in an adult sample. Journal of Personality, 54, 385-405.

Ogles, B.M., Lambert, M.J., & Craig, D.E. (1991). Comparison of self-help books for coping with loss: Expectations and Attributions. Journal of Counseling Psychology, 38, 387-393.

Orwin, R.G. (1983). A fail-safe N for effect size in meta-analysis. Journal of Educational Statistics, 8, 157-159.

Paul, G.L. (1969). Behavior modification research: Design and tactics. In C.M. Franks (Ed.), Behavior therapy: Appraisal and status. New York: McGraw-Hill.

Peterson-Delaney, S. (1938). Book therapy in veterans' hospitals. Library Journal, 63, 305-308.

Pomeroy, E. (1927). Book therapy in veterans' hospitals. In R.J. Rubin (Ed.), Bibliotherapy sourcebook (pp. 6-11). Phoenix, AZ: Oryx Press, 1978.

Posavac, E. & Carey, R.G. (1985). Program evaluation. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Reppucci, N.D., & Baker, B.L. (1969). Self-desensitization: Implications for treatment and teaching. In R.D. Rubin & C.M. Franks (Eds.), Advances in behavior therapy. New York: Academic Press.

Riordan, R.J. & Wilson, L.S. (1988). Bibliotherapy: Does it work? Journal of Counseling & Development, 67, 506-508.

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

Rogers, C. & Stevens, B. (1967). On becoming a person. New York: Simon & Schuster.

Rosen, G.M. (1978). Suggestions for an editorial policy on the review of self-help treatment books. Behavioral Therapy, 9, 90.

Rosen, G.M. (1981). Guidelines for the review of do-it-yourself treatment books. Contemporary Psychology, 26, 189-191.

Rosen, G.M. (1987). Self-help treatment books and the commercialization of psychotherapy. American Psychologist, 42, 46-51.

Rosenthal, R. (1991). Meta-analytic procedures for social research (rev. ed.) Newbury Park, CA: Sage Publications.

Rubin, R.J. (Ed.). (1978a). Bibliotherapy sourcebook. Phoenix: Oryx Press.

Rubin, R.J. (1978b). Using bibliotherapy: A guide to theory and practice. Phoenix: Oryx Press.

Saper, Z., & Forest, J.J. (1987). Personality variables and interest in self-help books. Psychological Reports, 60, 563-566.

Schallow, J.R. (1975). Locus of control and success at self-modification. Behavior Therapy, 6, 667-671.

Schrank, F.A., & Engels, D.W. (1981). Bibliotherapy as a counseling adjunct: Research findings. Personnel & Guidance Journal, 60, 143-147.

Schultheis, M. (1972). A guidebook for bibliotherapy. Glenview, IL: Psychotechnics, Inc.

Sclabassi, S.H. (1973) Literature as a therapeutic tool: A review of the literature on bibliotherapy. American Journal of Psychotherapy, 27, 70-77.

Scogin, F., Bynum, J., Stephens, G., & Calhoon, S. (1990). Efficacy of self-administered treatment programs: Meta-analytic review. Professional Psychology: Research and Practice, 21, 42-47.

Shadish, W.R., & Sweeney, R.B. (1991). Mediators and moderators in meta-analysis: There's a reason we don't let Dodo birds tell us which psychotherapies should have prizes. Journal of Consulting and Clinical Psychology, 59, 883-893.

Sheehy, G. (1974). Passages. New York: Bantam Books.

Smith, M. (1975). When I say no, I feel guilty. New York: Bantam Books.

Smith, D. & Burkhalter, J.K. (1987). The use of bibliotherapy in clincial practice. Journal of Mental Health Counseling, 9, 184-190.

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

Starker, S. (1988a). Psychologists and self-help books: Attitudes and prescriptive practices of clinicians. American Journal of Psychotherapy, 42, 448-455.

Starker, S. (1988b). Self-help treatment books: The rest of the story. American Psychologist, 43, 599-600.

Stevens, M.J. & Pfost, K.S. (1982). Bibliotherapy: Medicine for the soul? Psychology: A Quarterly Journal of Human Behavior, 19, 21-25.

Stuart,R.B., & Davis, B. (1972). Slim chance in a fat world: Behavioral control of obesity. Champaign, IL: Research Press.

Tews, R.M. (1970). Progress in bibliotherapy. In M.J. Voigt (Ed.), Advances in librarianship (Vol. 1). New York: Academic Press.

Tinsley, H.E.A. & Weiss, D.J. (1975). Interrater reliablity and agreement of subjective judgements. Journal of Counseling Psychology, 22, 358-376.

Wanous, J.P., Sullivan, S.E., & Malinak, J. (1989). The role of judgement calls in meta-analysis. Journal of Applied Psychology, 74, 259-264.

Wilkinson, L. (1992). SYSTAT for the MacIntosh, Version 5.2. [Computer program]. Evanston, IL: SYSTAT, Inc.

Zimbardo, P.G. (1977). Shyness. New York: Jove Publications.

VITA

Richard Welton Marrs was born on August 25, 1958, in Ada, Oklahoma. He is the son of Dick and Marilynn (Welton) Marrs of Arkansas City, Kansas. He has one sister, Diane (Marrs) Lawrence. He is married to Laura (Johns) Marrs and has two daughters, Brittany and Kaellyn.

He attended public schools in Arkansas City, Kansas and graduated from Arkansas City High School in 1976. His Associate of Arts degree was conferred in 1978 by St. John's College in Winfield, Kansas and his Bachelor of Arts degree was conferred in 1980 by Concordia (then College) University of River Forest, Illinois. In 1985 he obtained a Master of Science in Education with a major in counseling from the University of Kansas. He has also taken graduate coursework at Kansas State University and Concordia Seminary in St. Louis.

Rick Marrs has worked in agri-business, as a college admissions counselor and, since 1983, as a college instructor, professor and counselor.

He was accepted as a doctoral student in Counseling Psychology at Loyola University in 1986. His minor areas were career psychology and community psychology. His clinical internship was completed at the University Counseling Services of Kansas State University.

Since 1986 Rick has been working as an Assistant Professor of Psychology at Concordia University of River Forest, Illinois.

APPROVAL SHEET

The dissertation submitted by Richard Welton Marrs has been read and approved by the following committee:

Dr. Steven D. Brown, Director
Professor, Counseling and Educational Psychology
Loyola University of Chicago

Dr. Gloria Lewis
Associate Professor, Counseling and Educational Psychology
Loyola University of Chicago

Dr. Joseph Durlak
Professor, Psychology
Loyola University of Chicago

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

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

4/24/84
Date

Dr. Steven D. Brown, Director