Journal of Abnormal Child Psychology, Vol. 23, No. 1, 1995

Child and Adolescent Psychotherapy Outcomes in Experiments versus Clinics: Why the Disparity?

John R. Weisz,[1,2] Geri R. Donenberg,[1] Susan S. Han,[1] and Danika Kauneckis[1]

In a recent article, Weisz, Weiss, and Donenberg (1992) compared the effects of child and adolescent psychotherapy in experimental studies and in studies of clinic practice. Here we update that report with new information and we explore 10 possible reasons why, to date, therapy in experiments appears to have shown larger effect sizes than therapy in clinics. We find that beneficial therapy effects are associated with three factors which are more common in research therapy than in clinic therapy: (a) the use of behavioral (including cognitive–behavioral) methods, (b) reliance on specific, focused therapy methods rather than mixed and eclectic approaches, and (c) provision of structure (e.g., through treatment manuals) and monitoring (e.g., through review of therapy tapes) to foster adherence to treatment plans. These three factors all involve dimensions along which clinic procedures could be altered.

In a recent article, Weisz, Weiss, and Donenberg (1992) summarized evidence on effects of psychotherapy with children and adolescents (here referred to collectively as *children*). Two lines of evidence were surveyed: (a) experimental studies of child therapy effects, as summarized in meta-analyses, and (b) studies of the effects of conventional clinic-based psychotherapy. In the present paper we briefly review and update both lines of

Manuscript received in final form January 16, 1994.
The work reported here was supported in part through grants from the National Institute of Mental Health (R01 MH38240 and R03 MH34652) to the first author. For a more detailed treatment of the issues discussed here, and a more detailed description of the studies and meta-analyses reviewed, see the book, *Effects of Psychotherapy with Children and Adolescents* (Weisz & Weiss, 1993).
[1]University of California at Los Angeles, Los Angeles, California 90024.
[2]Address all correspondence regarding this article, including requests for reprints, to John Weisz, Department of Psychology, Franz Hall, UCLA, 405 Hilgard Avenue, Los Angeles, California 90024-1563.

0091-0627/95/0200-0083$07.50/0

evidence, noting that the experimental or "research therapy" studies show effects superior to those of the "clinic therapy" studies. We then explore 10 possible explanations for the superior effects of research therapy, sketching lessons that may be learned from therapy experiments and applied to clinic practice.

It is useful to describe our task in the context of numerous calls for specificity in outcome research. In a widely cited challenge, a quarter century ago, Paul (1967) urged attention to the outcome question, "*What* treatment, by *whom*, is most effective for *this* individual with *that* specific problem, under *which* set of circumstances?" (p. 111). More recently, Kazdin (1991) cited Paul's challenge and suggested, "The global question about treatment effects has been replaced by one that focuses on the effects of alternative treatments, as applied to specific clinical problems and patient samples and by different types of therapists" (p. 786). We agree with Paul and Kazdin that attention to such refined, specific, and comparative questions about psychotherapy effects is important. However, such work may need to be complemented by ongoing attention to our overall progress in the field. There is value, we suspect, in asking the following question: Given the array of child problems that need attention, the array of methods used to address child problems, and the array of clinicians who use these methods, is there evidence that, on average, our efforts to help are really helping? Moreover, we suggest that attempts to address this broad question are apt to lead quite naturally into more specific and comparative analyses — a process that will be illustrated in the present paper.

EXPERIMENTAL EVIDENCE ON CHILD PSYCHOTHERAPY EFFECTS

Experimental evidence on child and adolescent psychotherapy effects is perhaps best summarized in the form of meta-analytic reviews (see Mann, 1990; Smith, Glass, & Miller, 1980; but see also critiques of meta-analysis, e.g., by Wilson, 1985).

Broad-Based Meta-Analyses of Child Psychotherapy Research

We know of four broad-based child therapy meta-analyses, each focused on a broad array of types of intervention. Casey and Berman (1985) surveyed outcome studies published between 1952 and 1983, and focused on children aged 12 and younger. Mean effect size (ES) $\hat{w}$as 0.71 for the 64 studies that included treatment–control comparisons; the average

treated child functioned better after treatment than 76% of control group children, averaging across outcome measures. Weisz, Weisz, Alicke, and Klotz (1987) reviewed 105 outcome studies, published between 1952 and 1983, and including children aged 4 to 18. The mean ES was 0.79; the average treated child functioned better after treatment than 70% of control group peers. Kazdin, Bass, Ayers, and Rodgers (1990) surveyed studies published between 1970 and 1988, including youngsters aged 4 to 18. For 64 studies that compared treatment groups and *no-treatment control groups*, mean ES was 0.88, indicating that the average treated child was better off after treatment than 81% of the no-treatment youngsters. Some 41 studies in the Kazdin et al. collection involved treatment groups versus *active control groups*; for these studies, mean ES was 0.77, indicating that after treatment the average treated child was functioning better than 78% of the control group. Finally, we have preliminary data from a recent meta-analysis by Weisz, Weiss, Morton, Granger, and Han (1993), based on 110 studies published between 1967 and 1991, and involving children aged 2 to 18. Mean ES was 0.71, indicating that, after treatment, the average treated child was functioning better than 76% of control group children.

The evidence from these four broad-based meta-analyses indicates rather consistent positive effects; ES values ranged from 0.71 to 0.84 (estimated overall mean for Kazdin et al., 1990), near Cohen's (1988) threshold of 0.80 for a "large" effect. Table I summarizes findings of these four meta-analyses. The table includes information on (a) whether mean ES was found to be higher for behavioral than nonbehavioral interventions, (b) whether higher mean ES values were associated with treatment of females than males (no meta-analysis found males > females), (c) whether mean ES was higher for studies treating young subjects than for studies treating older ones (no meta-analysis found older > younger), (d) whether recent studies were found to show higher mean ES than older studies, and (e) whether ES was higher in studies focused on child overcontrolled (or internalizing) problems or in studies focused on undercontrolled (or externalizng) problems.

Specialty Meta-Analyses

Four more narrowly focused meta-analyses add to the corpus of available information on child psychotherapy effects. Hazelrigg, Cooper, and Borduin (1987) surveyed studies of the outcome of *family therapy.* Dush, Hirt, and Schroeder (1989) reviewed outcome studies that involved use of the cognitive–behavioral technique of *self-statement modification* (SSM) with children aged 5 to 16. Reviewing somewhat more broadly, Durlak, Fuhrman, and Lampman (1991) surveyed 64 outcome studies involving *cog-*

Table I. Overview of Four Broad-Based Meta-Analyses of Child Psychotherapy Research

	Casey-Berman (1985)	Weisz et al. (1987)	Kazdin et al. (1990)[a]	Weisz et al. (1992)
Number of studies	64	105	64/41	110
Mean age across studies	8.9	10.2	10.2	10.3
Age range (study means)	3-12	4-18	5-18	2-18
Percent male subjects	60	66	67	63
Effect size (ES) findings:				
Mean posttherapy ES	0.71	0.79	0.88/0.77	0.71
Percentile of treated subjects	76	79	81/78	76
Mean followup ES	0.60	0.93	0.89/0.39	0.61
Behavioral > nonbehavioral?	Yes/no[b]	Yes	NE	Yes
Females > males?	Yes	No	NE	Yes
Younger > older subjects?	No	Yes	NE	No
Recent > older studies?	No	NE	NE	No
Over- > undercontrolled?	NE	No	NE	No
ES formula used	M1-M2 SD (pooled)	M1-M2 SD (controls)	M1-M2 SD (pooled)	M1-M2 SD (controls)

[a]Kazdin et al. (1990) identified 223 outcome studies, but carried out separate meta-analyses for studies involving treatment versus no treatment comparisons (n = 64 studies) and studies involving treatment versus active control group comparisons (n = 41 studies). In the Kazdin et al. column, when we list one figure only, this refers to the full sample of 223; when we list two figures (e.g., 0.88/0.77), this refers to the two separate meta-analytic samples.
[b]An initial analysis showed behavioral > nonbehavioral, $p < 0.001$; an analysis controlling type of outcome measure reduced p to 0.06 (Casey, 1992). NE = Not examined in this meta-analysis.

nitive–behavioral therapy with children aged 4 to 13. Baer and Nietzel (1991) reviewed 36 outcome studies involving cognitive and behavioral treatment of child impulsivity. And finally, Russell, Greenwald, and Shirk (1991) explored the impact of child psychotherapy (of various types) on children's *language proficiency*. These five specialty meta-analyses all showed beneficial effects of psychotherapy.

LIMITATIONS OF THE EVIDENCE: RESEARCH THERAPY VERSUS CLINIC THERAPY

The nine meta-analyses noted above point to rather consistent positive effects of child psychotherapy. However, as noted elsewhere (e.g., Weisz & Weiss, 1993; Weisz et al., 1993), the 250-plus studies included in the meta-analyses constitute only part of the evidence we need on child psychotherapy outcome. Most of the studies included in the meta-analyses

(particularly the behavioral studies and studies of the past decade) appear to have involved children, interventions, and/or treatment conditions that are not very representative of conventional clinical practice. In many of these studies, (a) youngsters were recruited for treatment and were not actual clinic cases; (b) samples were selected for homogeneity, with therapy addressing one or two focal problems (e.g., a specific phobia); (c) therapists received concentrated pretherapy training in the specific intervention techniques they would use; and/or (d) the therapy involved primary or exclusive adherence to those specific techniques. In addition, (e) therapy was often highly structured, guided by a manual and/or monitored for its adherence to a treatment plan. (For other differences between psychotherapy research and therapy in clinical practice, see Kazdin, Bass, et al., 1990.) These features of the experimental studies tended to coalesce around a genre that might be termed *research therapy*, a genre that differs in a number of ways from conventional *clinic therapy*. Given the differences, a question arises as to whether the positive outcomes that have been demonstrated in the research therapy studies and summarized in the meta-analyses are replicated in actual clinical practice with children.

EVIDENCE ON THE EFFICACY OF CLINIC THERAPY

Answering this question is an important but difficult task. Clinic mandates almost always prevent randomly assigning those who seek treatment to no-treatment control conditions. Thus, in most cases, the virtue of random assignment is not available to those who seek to assess treatment effects in clinics. This being the case, researchers have tried to develop methods that do not require random assignment to control groups.

Levitt's (1957) Cross-Study Comparison

In an early effort in this direction, Levitt (1957b), following Eysenck (1952), attempted to estimate the effects of child therapy. From two studies of child clinic dropouts, Levitt estimated the base rate of improvement without treatment. He then compared this base rate to the improvement rate in 18 published reports of posttherapy outcome and 17 published reports of outcome at followup. The improvement rate was 72.5% for untreated children, 74% for treated children, a picture that offered no support for the efficacy of child therapy. Levitt's (1957b) report stimulated several critiques (e.g., Barrett, Hampe & Miller, 1978; Heinicke & Goldman, 1960; Hood-Williams, 1960),

some focused on the fact that Levitt had estimated base rates of improvement for treated and "untreated" groups from different studies.

Other Approaches

Other clinic-based studies have provided more direct comparisons of treated and untreated groups. We carried out a search for such studies, aiming for those that involved (1) treatment of clinic-referred (not recruited or "analog") youngsters, (2) treatment in ongoing, service-oriented clinics or clinical agencies, not in research settings (e.g., not a university lab or elementary school), (3) therapy conducted by professional clinicians (as opposed to, say, research assistants), and (4) therapy done as a part of the regular service-related program of the clinic, not primarily for research. We also required that the studies involve direct comparison between youngsters who received treatment and a control group who received none or a placebo condition. This ruled out, for example, treatment in hospital and residential settings where "control groups" had some intervention. Also excluded were studies involving only parent treatment or teacher training, and studies focused on children being treated for medical conditions.

A computer search proved difficult, because key words were hard to identify for such a specialized collection of studies. So, instead, we searched issue-by-issue, for the years 1972 to 1991, through 23 journals known to publish research on child therapy or mental health services. We also searched the reference lists of several reviews of child psychotherapy research and of the meta-analyses cited above. The search persuaded us that clinic studies which meet the criteria outlined above are very rare. We found only nine studies that seemed to fit, and most of these had been published years ago; the studies involved four rather different research strategies (For a more detailed examination of the strategies and the studies, see Weisz & Weiss, 1993).

STRATEGY 1: TREATED CHILDREN VERSUS MATCHED CHILDREN IN THE GENERAL POPULATION

One clinic-based outcome research strategy involves comparing children who are treated in clinics to demographically and clinically matched children drawn from the general population. In the only study we found that used this approach, Shepherd, Oppenheim, and Mitchell (1966) formed 50 pairs, aged 5 to 15, who were matched for demographic and

clinical characteristics (e.g., problem profile). Shepherd et al. compared the groups on adjustment about 2 years after initial assessment. In-home clinical interviews indicated that 63% of clinic cases had improved over the 2 years versus 61% of the nonclinic children.

STRATEGY 2: COMPARING TREATMENT CASES TO EVALUATION-ONLY CASES

A second approach is to compare outcomes among clinic-referred individuals who received treatment to outcomes for others who received only evaluation. Using this approach, Witmer and Keller (1942)[3] compared the adjustment of two groups of former clients at a child guidance clinic: 85 who had received treatment and 50 who had received only "diagnostic study." The two groups—similar in age, intelligence test scores, gender, health, economic status, and referral source, at the time of referral—were followed up 6 to 13 years after clinic contact. After home visit data and social agency information were combined, a rating of *successful* was given to 48% of the diagnostic cases versus 28% of the treated cases; corresponding figures for an *improved* category were 30% versus 26%, and for an *unimproved* category, 22% versus 40%. Overall, the untreated group had better long-term outcomes than the treated cases.

STRATEGY 3: COMPARING TREATED CASES TO DROPOUTS

A third approach involves comparing treated and untreated groups, both admitted to the same treatment facilities in the same time period, and both assigned to treatment, but with untreated cases consisting of those who drop out prior to treatment. Of course, one might argue that "dropouts" and "therapy completers" could differ demographically and/or clinically (see,

[3]A methodological limitation of this study (one that may be very difficult to avoid with Strategy 2) led us to exclude it from previous reviews (Weisz & Weiss, 1993; Weisz, Weiss, & Donenberg, 1991): It is unclear what determined which cases were designated for "diagnostic study" rather than treatment. The article noted that the designation was sometimes based on parents' willingness to have their child treated, sometimes on child adjustment or prognosis. It is also noted that "Delinquency was the only problem that characterized a proportionately higher number of the diagnostic cases . . . [and] more of these children were referred to the clinic by the court. Delinquency is often considered . . . to be prognostically unfavorable, so it might be argued that the diagnostic group's likelihood of favorable adjustment was thereby handicapped" (p. 79). It also appears that groups were not well matched on maternal and home characteristics. These concerns obviously becloud interpretation of the findings. However, to present a maximally comprehensive collection of the clinic study evidence, we included this study in the present review.

e.g., Hood-Williams, 1960). Yet, published studies examining a broad range of child demographic and clinical variables have generally revealed negligible differences between such groups (see Gould, Shaffer, & Kaplan, 1985; Levitt, 1957a; McAdoo & Roeske, 1973; Weisz & Weisz, 1993; Weisz, Weiss, & Langmeyer, 1987, 1989); moreover, it is possible to test for differences prior to comparing groups on outcome measures. Thus, to the extent that selection bias and group differences can be ruled out, children who drop out may be an acceptable control group for outcome research in circumstances where no randomly assigned control group can be formed.

Several studies have used this general approach. Such studies vary in their methodological rigor; we have been lenient with regard to such variations, accepting all studies that involve Strategy 3 unless they have methodological problems that clearly render them uninformative. In the earliest study of this genre that we know of, Lehrman, Sirluck, Black, and Glick (1949) focused on 3- to 20-year-olds seen at Jewish Board of Guardians (JBG) Child Guidance Clinics in New York. Some 196 cases received treatment; another 110 cases did not, even though it was recommended. One year after cases had been closed, the percentage classified as "success" was significantly higher for treated cases than for controls (50.5% vs. 31.8%), and "failure" ratings were slightly more common among control than treated cases (30.0% vs. 26.0%).

In a second relevant study, Levitt, Beiser, and Robertson (1959) studied youngsters seen at Chicago's Institute for Juvenile Research (mean age: about 10); 237 treated cases were compared to 93 untreated dropouts. About 5 years after clinic contact, Levitt et al. assessed outcomes on 26 variables, including psychological tests, child and parent ratings, clinician ratings, and life adjustment indicators. There were no treatment versus dropout group differences beyond chance expectancy.

In a third Strategy 3 study, Ashcraft (1971) followed up matched groups of treated (n = 40) and dropout (n = 43) cases who had been seen at one of two urban clinics at some time during grades 3 to 6. Because all children in both groups were "underachievers" who had "learning difficulties" (Ashcroft, 1971, p. 339), outcomes were assessed via standardized measures of academic achievement. Assessed 5 years postclinic contact, the treatment and control groups did not differ reliably on any measure.

In the fourth study using Strategy 3, Jacob, Magnusson, & Kemler (1972) compared 45 former child guidance clinic "remainers" (p. 140) and 42 "terminators" ("unilaterally left therapy before four treatment sessions" [p. 140]). At 1- and 2-year followups, the only direct outcome comparisons of the remainer and terminator groups showed no significant group differences.

In a more recent study, Weisz and Weiss (1989) compared 93 therapy completers (i.e., completed at least five sessions, terminated with concur-

rence of therapist) and 60 dropouts (i.e., no therapy, although it had been recommended) from nine outpatient mental health clinics. At 6 months and 1 year postintake, three adjustment measures were collected: Child Behavior Checklist parent and teacher reports, and parents' severity ratings on children's original referral problems. There was no significant group difference on any measure, at 6 months or at 1 year.

STRATEGY 4: RANDOM ASSIGNMENT

In spite of the obstacles to doing random assignment research in clinics, two teams of investigators have managed to carry out what appear to be true random assignment studies relevant to this review.

De Fries, Jenkins, and Williams (1964) studied 6- to 15-year-olds, all described as seriously disturbed, all having a psychiatric diagnosis, and all in foster care. Some 27 demographically and clinically matched child pairs were formed, with one member of each pair randomly assigned to receive conventional welfare department services, the other member assigned to psychotherapy plus enhanced services for the foster family. After therapy, outside judges classified each child as *improved*, *no change*, or *worsened*; there were no significant treatment-control differences. Although a second goal of intervention had been to prevent institutionalization, institutionalization actually occurred somewhat more often among the treated children than among controls.

In a second random-assignment study, Smyrnios and Kirkby (1993) assigned 30 clinic-referred children and their parents to either time-unlimited or time-limited (12 sessions) psychodynamic therapy, or to a minimal-contact control group. Outcome measures included goal attainment scales, parent ratings on target complaints, parent ratings on a family concept scale, and teacher reports of children's school behavior. At immediate post-clinic assessments, and at 4-year followup, some outcome measures showed no significant differences between treatment and control groups, and some pointed to somewhat better outcomes for the minimal contact control group than for the unlimited treatment group.

SUMMARIZING THE CLINIC-BASED FINDINGS

To summarize findings of the nine clinic-based studies for comparison with the meta-analytic evidence reviewed earlier,[4] we computed, for each

[4]In comparing ES values from the meta-analyses of predominantly lab studies with ES values from the clinic studies, one should bear in mind a common difference between lab and clinic

of the studies, an ES or an ES estimate (for reports that did not include the statistics needed for standard ES calculation, we used estimation procedures described by Smith et al., 1980, and Glass, McGaw, & Smith, 1981). For studies reporting results with more than one outcome measure and/or more than one point of outcome assessment (e.g., 1-year and 2-year followups), we computed separate ES values for each measure and/or each assessment point, then computed the mean of these values to generate a single ES value for each treatment group in each study.[5] As Table II shows, the 10 resulting ES values ranged from –0.40 to +0.29. The mean ES across studies (based on one ES per study) was 0.01.

LIMITATIONS AND IMPLICATIONS OF THE CLINIC-BASED FINDINGS

If the findings of these clinic studies were taken at face value, they would suggest that, for clinic-referred children treated by practicing clinicians in clinic settings, the impact of psychotherapy may not be as positive as the impact of research therapy has been in the experimental outcome

studies: Lab studies tend to involve a fixed number of therapy sessions, with outcome assessment occurring soon after the end of therapy, and sometimes again in a followup assessment. By contrast, the duration of clinic therapy tends to be different from one child to the next, and outcome assessment in clinic therapy studies is typically delayed until the last child in the sample has terminated (and sometimes much longer). As a result, immediate posttreatment outcome assessment is quite rare, and delayed outcome assessment is quite common, in clinic therapy studies. Thus, the ES values generated for clinic studies tend to reflect outcomes assessed well after treatment has been completed, whereas the ES values reported in meta-analyses tend to involve immediate posttreatment outcomes and relatively short-term followup assessments. However, it is not clear that lag time between treatment and outcome assessment has a major impact on ES values. Our meta-analyses (Weisz et al., 1987, 1993) have found nonsignificant differences between posttreatment and followup ESs; and the clinic (see Table II) and inpatient (see Tale III) studies we have identified also show negligible differences between earlier and later followups (with a slight trend toward higher ES for later followups).

[5]Three studies (noted in Table II) used shorter-term and longer-term outcome assessments; our decision to average across the two different assessment points for these studies bears examination. At first blush, it might seem most appropriate to rely only on the shorter-term assessment, to maximize comparability between clinic study ESs and the relatively short-term lab study ESs reflected in the meta-analyses reported earlier (see footnote 4). We chose to average ES values across the two assessments in these three clinic studies, however, because it was not clear which assessment would be the fairest representation of treatment outcome. Some believe that therapy effects are strongest near the end of treatment, but others believe that effects build over time, and this would make later assessment more appropriate. Because later followups tended to show slightly higher mean ES than did earlier followups, our decision to average across outcome assessment points had the effect of raising the mean ES for clinic studies slightly above the mean that would have resulted from use of only the shorter-term assessment in each study. This is also true of our ES assessment for "add-on studies," shown in Table III.

Table II. Effect Size (ES) Estimates for the Clinic-Based Studies[a]

Study		ES
Ashcraft (1971) (Completer vs. dropout)		0
De Fries, Jenkins, & Williams (1964) (Random assignment)		−0.12
Jacob, Magnussen, & Kemler (1972) (Completer vs. dropout)		0.09[b]
Lehrman, Sirluck, Black, & Glick (1949) (Completer vs. dropout)		0.28
Levitt, Beiser, & Robertson (1959) (Completer vs. dropout)		−0.05[c]
Shepherd, Oppenheim, & Mitchell (1966) (Therapy vs. general population)		−0.01
Smyrnios & Kirkby (1993) (Random assignment)	Time-limited versus control	0.29[b]
	Time-unlimited versus control	−0.10[b]
Weisz & Weiss (1989) (Completer vs. dropout)		0.18[b]
Witmer & Keller (1942) (Therapy vs. evaluation only)		−0.40

[a]The research strategy used to assess therapy effects is shown within the parentheses beside each study.
[b]Mean ES based on an average of shorter-term and longer-term outcome assessments.
[c]This mean ES is based on 16 of the 26 outcome measures reported by Levitt et al. Ten of the 26 were judged inappropriate for our calculations, most because they were only relevant to a subset of the sample (e.g., military record, which was not relevant to females or males who had not served).

studies. On the other hand, the nine clinic studies we have found represent a very small data base, much of it rather dated. Clearly, we need to build a base of careful outcome research on therapy as *currently* practiced in clinics, research that meets current standards of methodological rigor. Even if such research should replicate the modest findings of the clinic studies reported here, such an outcome would not need to be seen as entirely discouraging. The good news is that the research therapy studies, summarized in the meta-analyses, indicate that under the right conditions child therapy may be very effective. An important research objective thus becomes that of identifying those "right conditions" under which effects of child therapy may be optimized. One step in this process is the identification of factors that may account for the superior effects of research therapy over clinic therapy (assuming that the findings presented above hold up in future research).

WHY ONE EFFECT SIZE DOES NOT FIT ALL: EXPLORING 10 HYPOTHESES

This brings us to the second objective of the present paper. As implied above, there are several possible reasons why the mean ES values found in the meta-analyses have not been replicated in the clinic studies reviewed here. Here we will examine some of the possibilities in a preliminary way. To do so, we describe relevant analyses from previous meta-analytic reviews, and we report new analyses of a collection of child and adolescent treatment outcome studies assembled for our 1993 meta-analysis described above (Weisz et al., 1993). In essence, we test possible explanations of the research therapy versus clinic therapy outcome difference by carrying out focused analyses of relevant variables in a large research therapy data set. Clinic studies are obviously too scarce to be used in this way, and it would be inappropriate to mix research therapy and clinic therapy studies, because the research versus clinic distinction would be partially confounded with the explanatory variables of interest. The 1993 collection of studies includes articles published between 1967 and 1991. We excluded studies which were not directly relevant (e.g., because their therapy involved mediated treatment — e.g., parent training or teacher training). The resulting pool included 64 studies, involving 100 treatment–control comparisons. We must stress that the analyses we report here are preliminary. The collection of studies which we employ here is certainly not complete, and it is possible that the samples are biased in ways that we do not recognize (e.g., because of the particular journals we have searched and those we have overlooked). Also, each group difference (i.e., in mean ES) that we report below involves comparison of groups which differed in their position on one particular variable (e.g., clinic-referred vs. analog subjects); as in meta-analyses generally, the variable of interest in any particular test may be confounded with a number of other potentially relevant variables, and this hampers clear interpretation of the findings. Despite these caveats, we think it is useful to begin exploring possible reasons for the clinic therapy versus research therapy difference, emphasizing, as we do, that the preliminary analyses we report here are best regarded as hypothesis-generating exercises. Altogether, we consider 10 possible explanations of why outcomes reported for research therapy are generally superior to outcomes reported for clinic therapy.

1. *Most of the clinic studies available in the literature were published years ago, when therapy methods and assessment technology were relatively primitive; the research therapy studies have better outcomes because they are more recent and thus reflect better methodology.* The clinic therapy studies have a relatively early mean publication date. They range from 1942 to

1993, but seven of the nine studies were published more than 20 years ago; the research therapy studies summarized in the meta-analyses have considerably more recent publication dates. If our methods of therapy — or our methods of assessing child mental health or therapy outcome — have improved over the years, then more recent studies would be expected to show larger effect sizes than older studies. This trend alone might account for the superior effects of research therapy.

To assess this possibility, we computed the correlation between ES and year of publication across the 64 studies from our 1993 data set. Casey and Berman (1985) also computed this correlation for the studies in their 1985 meta-analysis of child outcome studies (described above). Both calculations revealed a negligible relationship between ES and publication year. At first blush, this appears to indicate no significant improvement in effectiveness of therapy (or precision of assessment) over the years. This may be the case, but there is another possibility: It may be that child psychotherapy has grown more effective over the years, but that journals are increasingly likely to accept studies reporting null findings; these two trends might offset one another, thus wiping out any evidence of a secular trend in mean ES. Whatever the proper explanation for the absence of a trend, it is clear that the evidence we can identify provides no support for the "year of study" hypothesis as an explanation of the research therapy versus clinic therapy difference in outcomes.

2. *Research therapy studies have better outcomes because they are more likely than clinic studies to involve behavioral treatment methods.* The report by Kazdin, Bass, et al. (1990) on characteristics of therapy research indicates that the majority of treatment outcome research involves behavior modification (50% of all studies surveyed) and cognitive–behavioral interventions (22%), with scant attention given to psychodynamic, family, or eclectic interventions. Yet Kazdin, Siegel, and Bass's (1990) survey of child clinicians indicated that a majority viewed psychodynamic, family, and eclectic interventions as useful most or all of the time. Certainly it is true that behavioral (including behavior modification and cognitive–behavioral) interventions were not well-represented among the clinic studies reviewed above, but extremely well-represented in each of the four broad-based meta-analyses.

This pattern becomes potentially important when we examine the evidence on relative efficacy of behavioral and nonbehavioral methods. Three of the four child meta-analyses (i.e., Casey & Berman, 1985; Weisz & Weiss, 1993; Weisz et al., 1987, 1993) have included a comparison of the mean ES obtained for behavioral and nonbehavioral studies. In all three, the initial analysis showed higher mean ES for behavioral than nonbehavioral interventions (see Table I). There is certainly an ongoing debate about

the proper interpretation of such findings and about which variables should be controlled in a comparative analysis (see Casey & Berman, 1985; Weiss & Weisz, in press; Weisz et al., 1987). However, at present, the bulk of the evidence seems to point to higher ES values for behavioral than nonbehavioral interventions when one pools the available research.

If effect sizes are generally higher for behavioral than for nonbehavioral interventions, and if behavioral methods are more common in research therapy than in clinic therapy, then the superior effects of research therapy might be attributable to differences in methods of intervention. Meta-analytic findings appear to support this possibility.

3. *Clinic cases are more seriously disturbed, come from more dysfunctional families, or for some other reason are more difficult to treat successfully than youngsters recruited for research therapy.* The third possibility we examined is that the youngsters often seen in research therapy — i.e., recruited subjects often selected from schools — are, for some reason, easier to treat successfully than are typical clinic cases. Outcome research has been criticized for excessive reliance on recruited *analog* samples rather than on samples who were actually referred to clinics (see, e.g., Parloff, 1980; Shapiro & Shapiro, 1982). It is certainly possible that clinic-referred children are more seriously disturbed, that they come from more dysfunctional families, or that they are more likely to confront stressful life circumstances (e.g., low income levels) than are the often-middle-class youngsters whose parents give consent for their participation in research therapy studies.

We addressed this analog sample issue by comparing the mean ES achieved in studies using analog samples (i.e., subjects were recruited and nonreferred) to the mean ES in studies using clinical samples (i.e., in which subjects were clinic-referred or would have warranted treatment regardless of the study). This comparison was conducted in both our first meta-analysis (Weisz et al., 1987, p. 547) and our more recent one (Weisz et al., 1993). In the 1987 meta-analysis, we identified 126 treatment–control comparisons involving analog samples and 37 comparisons involving true clinical samples. The mean ES was 0.76 for the analog samples and 0.89 for the clinical samples, a nonsignificant difference running counter to the explanation being tested. In our more recent meta-analysis (Weisz et al., 1993), we found 129 treatment–control comparisons involving analog samples and 45 comparisons involving true clinical samples. The mean ES was 0.77 for the analog samples and 0.55 for the clinical samples, again a nonsignificant difference ($p > 0.40$). So the evidence did not support the hypothesis that research therapy studies achieve superior effect sizes because of differences in the treatability of analog and clinical subjects.

4. *There is something about clinical settings that undermines therapy effectiveness.* Another possibility is that some aspect of the clinic setting in-

terferes with the efficacy of therapy. For example, because clinics typically serve large numbers of children daily, youngsters and parents who visit clinics may be treated in a more routine manner than are the small number who are courted as "research subjects," sometimes paid for their participation, and treated in a solicitous manner to ensure that they attend sessions faithfully. Receiving one's therapy in clinics may involve unwanted time on a waiting list, completion of multiple medical and insurance forms, and intrusive questions about family income — irritations that are rarely a part of therapy in labs or school settings.

It is possible that these or other factors hamper the efforts of therapists who do their work in clinics. To explore this possibility, we distinguished, in the 1993 meta-analytic data set, between therapy that took place in clinical and nonclinical settings. For this variable, and others noted below, two coders independently classified 10 studies, involving 22 treatment–control comparisons. The treatment location variable included several categories: (a) research lab (usually in a university), (b) primary or secondary school, (c) outpatient clinic, (d) hospital, (e) psychiatric inpatient clinic, and (f) detention or correctional facility. The judges achieved a kappa of .78 on these judgments. Ratings by the one judge who coded all the studies were used in these analyses and others reported below.

We conducted a *t*-test comparing mean ES in clinical settings [items (c), (d), and (e) above; $n = 15$ treatment–control comparisons] versus nonclinical settings [(a) and (b) above; $n = 76$]. Mean ES was 0.53 for groups treated in clinical settings and 0.73 for groups treated in nonclinical settings. The *t*-test showed that the difference did not approach significance ($p > .30$). Thus, we found no reliable support for the notion that clinical settings are associated with less beneficial therapy effects than nonclinical settings.

5. *Clinicians (for reasons such as heavy caseloads) are less effective than research therapists.* Therapists in clinics are often burdened by heavy caseloads, extensive paperwork responsibilities, and random or "turn-taking" procedures for case assignment. Moreover, clinic therapists may have less opportunity than research therapists to receive fresh information and training in therapy methods. We explored the impact of the therapist variable on therapy outcome by coding studies in the 1993 meta-analytic data set for whether the therapists involved were *clinicians* (e.g., clinic employees, Masters of Social Work, marriage and family counselors, interns, school counselors), *researchers* (e.g., university faculty, study authors, graduate students in a Ph.D. program), or *both* (i.e., one or more therapists from both categories were involved in the therapy). The two raters obtained a kappa of .86 for these judgments. Mean ES was (coincidentally) 0.86 for the clinicians ($n = 29$) and 0.63 for the researchers ($n = 44$); the difference

was not significant ($p > .35$). Thus, the analysis offered no support for the notion that researchers were more successful therapists than clinicians.

6. *Research therapy is more effective because the therapists have special training, just prior to the intervention, in the methods that will be used.* A sixth possibility is that research therapy is more effective because research therapists, more often than clinic therapists, are prepared for their work via pretreatment training in the method of intervention that will be used. We classified treatment–control comparisons in the 1993 data set (again excluding studies involving mediated therapy) for whether the study (a) explicitly mentioned such pretreatment training, (b) implied such training, (c) explicitly stated that there was no such training, or (d) implied that there was no such training. The two raters achieved a kappa of .72 in these judgments. For studies involving pretreatment training [i.e., (a) and (b) above; $n = 45$] the mean ES was 0.47; for studies not providing such training [i.e., (c) and (d); $n = 11$], the mean ES was 0.70. The direction was counterintuitive, but the *t*-test revealed that the difference was not reliable ($p > .40$). so we found no support for the idea that special pretreatment training might underlie the superior effects of research therapy.

7. *Research therapy is more effective because it is more highly structured than clinic therapy.* One apparent difference between research therapy and clinic therapy is that the former tends to be more highly structured; research therapy often entails a preplanned set of procedures, used in a particular order, sometimes guided by a treatment manual, and quite often monitored for fidelity via videotapes, audiotapes, and/or direct observation of treatment sessions. Such structure may certainly reduce the flexibility of the therapy, limiting the therapist's ability to tailor procedures to the perceived needs of a particular child. However, relative inflexibility notwithstanding, adherence to a carefully planned treatment program may also enhance effectiveness by maximizing consistency in the implementation of the planed intervention. Whether structure and consistency enhance or undermine treatment efficacy is an empirical question, one which we addressed by classifying studies into several different groups: (a) highly structured (i.e., both a treatment manual and monitoring of therapist consistency via video, audio, direct observation, or review of therapist notes), (b) structured (i.e., either a treatment manual or monitoring), (c) semistructured (i.e., treatment outline used to guide therapists), (d) tape therapy (i.e., therapy was administered solely or primarily by the playing of a tape for subjects to view and/or listen to), and (e) unstructured (i.e., none of the above). The two independent coders achieved a kappa of .68 for these judgments.

To test the hypothesis, we compared structured therapy studies [categories (a), (b), (c), and (d) combined; $n = 82$] to unstructured therapy studies [category (e); $n = 16$], ES means were 0.72 for structured therapy,

0.39 for unstructured. The group difference was statistically significant, t (two-tailed, unequal variances, $df = 72.5$) = 2.504, $p = 0.01$. This result is consistent with the notion that structured therapy is more effective than unstructured approaches and, by implication, that the structured nature of research therapy contributes to its superior effects.

8. *Research therapy is more effective because it tends to involve a focus on one type of child or one type of child problem, whereas clinic therapy tends to be more diffuse*. It is part of the mandate in most clinics that every youngster who has legitimate mental health needs must be seen, at least to the extent that clinic resources permit. Clinics that serve the public rarely have the luxury of deciding to serve, say, only depressed children, or only those with an attention deficit with hyperactivity disorder (ADHD) diagnosis. Moreover, family concerns often dictate an array of problems needing attention for a given child, and clinic traditions often dictate that many or all of these problems are concurrently a focus of the intervention for that child.

By contrast, a readily identifiable feature of research therapy is targeting, or focus. Researchers typically select a rather narrow intervention target — e.g., a particular syndrome (e.g., ADHD), a constellation of problems centering on one theme (e.g., aggressive behavior), or even a single problem (e.g., out-of-seat behavior). It is possible that this narrow focus makes therapeutic success more likely, because the therapeutic objective is clearer and simpler, and the relevant treatment(s) easier to identify, than in the case of clinic therapy.

To explore this possibility, we used the *target problem* coding scheme from Weisz et al. (1987), focusing on the part ("tier 2") that classifies treated problems at a moderate level of specificity — e.g., "phobia" rather than "overcontrolled" or "fear of the dark" (intercoder kappa = .86). For the present analysis, studies were considered to have a *clear problem focus* if target problem had been coded delinquency, noncompliance, self-control (hyperactivity/impulsivity), aggression, phobias/anxiety, social withdrawal/isolation, depression, somatic problem (headaches), or underachievement. Studies were classified as having a *vague or diffuse problem focus* if target problem had been coded adjustment/emotional disturbance, multiple/vague problems, multiple/vague undercontrolled, or multiple/vague overcontrolled. A t-test comparing those treatment–control comparisons involving a clear problem focus ($n = 66$) to the comparisons involving a vague or diffuse focus ($n = 33$) was not significant ($p > .20$; mean ES was, counterintuitively, 0.58 for the *clear* group and 0.86 for the *diffuse* group). Thus, we found no support for the notion that precise problem focus might explain the superior effects of research therapy.

9. *Research therapy is more successful because it tends to involve more specific, focused treatment methods than clinic therapy.* Partly as a result of the diverse array of children and problems clinic therapists see, these therapists may find themselves using an array of therapeutic methods rather than being doctrinaire. Indeed, when Kazdin, Siegel, and Bass (1990) surveyed 1162 child and adolescent clinicians and asked which treatment approaches they considered "useful most or all of the time," the most commonly endorsed approach was "eclectic" (endorsed by 72% of the sample). It is possible that therapists in clinics are more likely than research therapists to draw rather eclectically from an armamentarium of approaches during the treatment of a given child. Research therapists appear more likely to use a single, focused approach to therapy, often one that fits a particular theoretical model rather precisely.

It is possible that the superiority of research therapy over clinic therapy results in part from the more focused therapy methods often used in research studies. To explore this possibility, we focused on codes within the *therapy type* coding scheme of Weisz et al. (1987). We used two levels, or "tiers" of that coding scheme. Tier 2 (kappa = .71) classifies therapy methods into the following categories: operant, respondent, modeling, social skills training, cognitive/cognitive–behavioral, multiple behavioral, client-centered/nondirective, insight-oriented/psychodynamic, and discussion group. Tier 3 (kappa = .78) classifies therapy methods into rather specific subtypes grouped beneath the tier 2 categories (e.g., under the tier 2 category of operant methods are listed such tier 3 categories as physical reinforcement and social/verbal reinforcement). We classified treatment groups in the 1993 data set (excluding mediated therapy studies) as involving *specific, focused treatment* if they could be coded into one of the most specific tier 3 categories (e.g., physical reinforcement, social/verbal reinforcement, systematic desensitization), *moderately focused treatment* if they fit a general category only (e.g., social skills training, cognitive/cognitive–behavioral, multiple modeling methods), and *vague or diffuse treatment* if they had been coded into the tier 2 categories of discussion group, mixed behavioral and nonbehavioral, or too vague to be classified. We then compared the specific, focused treatment studies (n = 61 treatment–control comparisons) to the vague or diffuse treatment studies (n = 6 comparisons). Mean ES was 0.77 for the specific, focused group and 0.32 for the vague or diffuse group, t (two-tailed, unequal variances, df = 36.5) = 3.13, $p < 0.01$. The analysis thus supported the possibility that specific, focused treatment may have contributed to the superior outcomes of research therapy.

10. *The research therapy versus clinic therapy outcome difference results in part from the differential length of therapy in the two conditions.* The final possibility we examined was that there is something about the differential

length of typical research therapy versus typical clinic therapy that accounts for the superior outcomes of research therapy. Kazdin, Bass, et al. (1990) noted that outcome studies typically involve brief interventions on the order of 8 to 10 weeks, whereas clinical practitioners surveyed by Kazdin, Siegel, and Bass (1990) reported a mean treatment duration of 27 weeks (25 for psychologists, 35 for psychiatrists). Although intuition might suggest that more treatment is better than less, it is also possible that brief, time-limited psychotherapy imposes a certain discipline on the therapist, encouraging attentiveness to the tasks of therapy and efficiency in the use of therapeutic techniques. Conceivably, then, the brevity of the process may enhance the effectiveness of research therapy. To explore this issue, we computed correlations between ES and length of therapy in each of the studies in the 1993 data set. Using number of therapy hours as the measure of duration, we found a nonsignificant correlation of –0.12 between duration and ES. This result obviously provides no support for a relationship between therapy duration and therapy effect, at least within the limits of the data reviewed here.

COMPARING THE CLINIC STUDIES TO INPATIENT AND RESIDENTIAL ADD-ON STUDIES

As a complement to tests of the 10 specific explanations, above, we examined 10 studies of the outcome of special treatment interventions that had been added to the standard program of various inpatient or residential facilities. In each study, one or more treatment groups were compared to one or more control groups of youth in the same facility who did not receive the special treatment intervention. The studies were identified through our literature search for clinic studies, as described above; we excluded studies in juvenile detention and correctional facilities. The ten studies are briefly described in Table III.

We calculated an ES (or estimate) for each treatment group in each study. Across the 10 studies, with one ES per study (averaging across posttreatment and followup assessments), we found a mean ES of 0.76; this indicated that the average youngster receiving one of the special interventions was, after treatment, at the 78th percentile of the control group, averaging across measures of adjustment. This ES value was significantly different from 0, $t(9) = 3.43$, $p < 0.005$, and quite similar to the mean ES of the four meta-analyses in Table I. On the other hand, this mean ES of 0.76 was markedly different from the clinic study mean ES of 0.01, t (two-tailed, unequal variances, $df = 10.6) = 3.25$, $p < 0.01$.

This difference is informative, given the ways these add-on studies resemble and differ from the clinic studies. The add-ons resemble the clinic studies in that (1) both involve samples not of recruited children but of youngsters seriously enough disturbed to have been referred for treatment (indeed, the inpatient and residential samples may well be *more* disturbed than the clinic samples); (2) both took place in clinical settings, and (3) both primarily involve therapy conducted by clinicians. These three similarities are consistent with the findings of the preceding analyses — i.e., consistent with the notion that the relatively low ES in clinic studies is *not* attributable to the severity of the young clients' problems, the clinical settings in which the treatment takes place, or the fact that therapists are employed as clinicians.

By contrast, the add-on studies *differ from* the clinic studies in ways consistent with findings of our clinic study analyses. Specifically, the add-on studies, more often than the clinic studies (1) used predominantly behavioral (including cognitive–behavioral) interventions, (2) used rather precisely focused intervention methods, and (3) involved structured therapy methods with monitoring to insure therapist consistency and fidelity to the treatment plan. Thus, both lines of analysis employed here leave open the possibility that these three factors may help to account for the superior effects of research therapy over clinic therapy.

CONCLUDING COMMENT

We examined 10 possible explanations for the research therapy versus clinic therapy difference in measured effectiveness. Three candidate explanations found support in our analyses of the general meta-analytic data. Impressions from those analyses were consistent with what we found when we compared clinic studies to add-on studies in hospital and residential settings. It would be a mistake, however, to construe the present analyses as definitive. What we have done here is generate some potentially useful hypotheses. These need to be tested in more precise ways in future research.

If future research should support these hypotheses, one result might be a reconsideration of how child treatment is conducted in many clinics. It may be important that some of the factors examined here that would have been most difficult for clinics to alter (e.g., clinic settings, the fact that clinic therapy is done by clinic-employed professionals, the fact that children seen in clinics may be seriously disturbed and/or referred for multiple, diverse problems) were *not* supported as relevant explanatory factors in the analyses. Instead, those explanatory factors supported in the present

Table III. Effect Size (ES) Estimates for the Add-On Studies in Hospital and Residential Settings

Study	ES
Finch, Wilkinson, Nelson, & Montgomery (1975) (Self-instruction to modify boys' cognitive tempo)	2.62[a]
Jennings & Davis (1977) (Structured learning of social skills)	0.29[b]
Kazdin, Esveldt-Dawson, French, & Uni (1987) (Cognitive-behavioral problem solving for antisocial)	0.96[b]
(Nondirective relationship therapy for antisocial)	0.17[b]
Kendall & Finch (1978, 1979) (Cognitive-behavioral for impulsivity)	0.44[b]
Kolko, Loar, & Sturnick (1990) (Social-cognitive skills training for disruptive behavior)	0.86[b]
Krop, Calhoon, & Verrier (1971) (Covert reinforcement to modify self-concept)	0.86[b]
(Overt reinforcement to modify self-concept)	0[b]
Moran, Fonagy, Kurtz, Bolton, & Brook (1991) (Psychoanalytic treatment with brittle diabetics)	0.85[b]
Nahme-Huang, Singer, Singer, & Wheaton (1977) (Imaginative play for disturbed children)	0.16
(Perceptual–motor intervention for disturbed children)	0.15
Reardon & Tosi (1977) (Cognitive restructuring, relaxation, etc., adolescent girls in residential treatment for delinquency)	0.56[a,b,c]
Snyder & White (1979) (Cognitive self-instruction for externalizing)	1.56[b]
(Contingency awareness training for externalizing)	0.02[b]

[a]Mean ES value averaged across comparisons with multiple control groups.
[b]Mean ES based on an average of shorter-term and longer-term outcome assessments.
[c]This study included two treatment conditions: (a) rational stage directed therapy, and (b) rational restructuring with no relaxation. Our ES figure applies only to the first of these treatments; the article did not provide sufficient information on outcome for treatment b to permit an ES calculation.

analyses all reflect dimensions along which change is possible. It would certainly be possible, for example, for clinics to move toward heavier reliance on behavioral interventions, to emphasize focused and specific forms of therapy, and to ensure that therapy is structured (e.g., via manuals) and monitored for fidelity and consistency (e.g., through review of taped sessions). Indeed, the overall picture of what may be needed is analogous to what the gardener does to water a particularly distant, difficult-to-reach spot — i.e., shifting the hose from the "spray" to the "stream" setting. It

would be premature, through, to recommend such changes before more thorough and carefully controlled analyses can be brought to bear on the potential explanations addressed here. Our hope is that the ideas offered in this paper will stimulate such analyses.

REFERENCES

Ashcraft, C. W. (1971). The later school adjustment of treated and untreated emotionally handicapped children. *Journal of School Psychology, 9*, 338-342.

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

Barrett, C. L., Hampe, I. E., & Miller, M. C. (1978). Research on child psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (2nd ed., pp. 411-435). New York: Wiley.

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

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

De Fries, Z., Jenkins, S., & Williams, E. C. (1964). Treatment of disturbed children in foster care. *American Journal of Orthopsychiatry, 34*, 615-624.

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

Dush, D. M., Hirt, M. L., & Schroeder, H. E. (1989). Self-statement modification in the treatment of child behavior disorders: A meta-analysis. *Psychological Bulletin, 106*, 97-106.

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16*, 319-324.

Finch, A. J., Jr., Wilkinson, M. D., Nelson, W. M. III, & Montgomery, L. E. (1975). Modification of an impulsive cognitive tempo in emotionally disturbed boys. *Journal of Abnormal Child Psychology, 3*, 49-52.

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

Gould, M. S., Shaffer, D., & Kaplan, D. (1985). The characteristics of dropouts from a child psychiatry clinic. *Journal of the American Academy of Child Psychiatry, 24*, 316-328.

Hazelrigg, M. D., Cooper, H. M., & Borduin, C. M. (1987). Evaluating the effectiveness of family therapies: An integrative review and analysis. *Psychological Bulletin, 101*, 428-442.

Heinicke, C. M., & Goldman, A. (1960). Research on psychotherapy with children: A review and suggestions for further study. *American Journal of Orthopsychiatry, 30*, 483-494.

Hood-Williams, J. (1960). The results of psychotherapy with children: A revaluation. *Journal of Consulting and Clinical Psychology, 24*, 84-88.

Jacob, T., Magnussen, M. G., & Kemler, W. M. (1972). A follow-up of treatment terminators and remainers with short-term and long-term symptom duration. *Psychotherapy: Theory, Research, and Practice, 9*, 139-142.

Jennings, R. L., & Davis, C. S. (1977). Attraction-enhancing client behaviors: A structured learning approach for "non-YAVIS, Jr." *Journal of Consulting and Clinical Psychology, 45*, 135-144.

Kazdin, A. E. (1991). Effectiveness of psychotherapy with children and adolescents. *Journal of Consulting and Clinical Psychology, 59*, 785-798.

Kazdin, A. E., Bass, D., Ayers, W. A., & Rodgers, A. (1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting and Clinical Psychology, 58*, 729-740.

Kazdin, A. E., Esveldt-Dawson, K., French, N. H., & Unis, A. S. (1987). Problem-solving skills training and relationship therapy in the treatment of antisocial child behavior. *Journal of Consulting and Clinical Psychology, 55*, 76-85.

Kazdin, A. E., Siegel, T. C., & Bass, D. (1990). Drawing on clinical practice to inform research on child and adolescent psychotherapy: Survey of practitioners. *Professional Psychology: Research and Practice, 21*, 189-198.

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

Kendall, P. C., & Finch, A. J., Jr. (1979). Analysis of changes in verbal behavior following a cognitive-behavioral treatment for impulsivity. *Journal of Abnormal Child Psychology, 7*, 455-463.

Kolko, D. J., Loar, L. L., & Sturnick, D. (1990). Inpatient social-cognitive skills training groups with conduct disordered and attention deficit disordered children. *Journal of Child Psychology and Psychiatry, 31*, 737-748.

Krop, H., Calhoon, B., & Verrier, R. (1971). Modification of the "self-concept" of emotionally disturbed children by covert reinforcement. *Behavior Therapy, 2*, 201-204.

Lehrman, L. J., Sirluck, H., Black, B. J., & Glick, S. J. (1949). *Success and failure of treatment of children in the child guidance clinics of the Jewish Board of Guardians, New York City* (Research Monographs No. 1). New York: Jewish Board of Guardians.

Levitt, E. E. (1957a). A comparison of "remainers" and "defectors" among child clinic patients. *Journal of Consulting Psychology, 21*, 316.

Levitt, E. E. (1957b). The results of psychotherapy with children: An evaluation. *Journal of Consulting Psychology, 21*, 189-196.

Levitt, E. E., Beiser, H. R., & Robertson, R. E. (1959). A follow-up evaluation of cases treated at a community child guidance clinic. *American Journal of Orthopsychiatry, 29*, 337-347.

Mann, C. (1990). Meta-analysis in the breech. *Science, 249*, 476-480.

McAdoo, W. G., & Roeske, N. A. (1973). A comparison of defectors and continuers in a child guidance clinic. *Journal of Consulting and Clinical Psychology, 40*, 328-334.

Moran, G., Fonagy, P., Kurtz, A., Bolton, A., & Brook, C. (1991). A controlled study of the psychoanalytic treatment of brittle diabetes. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*, 926-935.

Nahme-Huang, L., Singer, D. G., Singer, J. L., & Wheaton, A. B. (1977). Imaginative play training and perceptual-motor intervention with emotionally disturbed hospitalized children. *American Journal of Orthopsychiatry, 47*, 238-249.

Parloff, M. B. (1980). Psychotherapy and research: An anaclitic depression. *Psychiatry, 43*, 279-293.

Paul, G. (1967). Outcome research in psychotherapy. *Journal of Consulting Psychology, 31*, 109-118.

Reardon, J. P., & Tosi, D. J. (1977). The effects of rational stage directed imagery on self-concept and reduction of psychological stress in adolescent delinquent females. *Journal of Consulting Psychology, 33*, 1084-1092.

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

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

Shepherd, M., Oppenheim, A. N., & Mitchell, S. (1966). Childhood behavior disorders and the child-guidance clinic: An epidemiological study. *Journal of Child Psychology and Psychiatry, 7*, 39-52.

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

Smyrnios, K., & Kirkby, K. X. (1993). Long-term comparison of brief versus unlimited treatments with children and their parents. *Journal of Consulting and Clinical Psychology, 61*, 1020-1027.

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

Weiss, B., & Weisz, J. R. (in press). Relative effectiveness of behavioral and nonbehavioral therapies with children and adolescents. *Journal of Consulting and Clinical Psychology.*

Weisz, J. R., & Weiss, B. (1989). Assessing the effects of clinic-based psychotherapy with children and adolescents. *Journal of Consulting and Clinical Psychology, 57,* 741-746.

Weisz, J. R., & Weiss, B. (1993). *Effects of psychotherapy with children and adolescents.* New York: Sage.

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

Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 47,* 1578-1585.

Weisz, J. R., Weiss, B., & Langmeyer, D. B. (1987). Giving up on child psychotherapy: Who drops out? *Journal of Consulting and Clinical Psychology, 55,* 916-918.

Weisz, J. R., Weiss, B., & Langmeyer, D. B. (1989). On "dropouts" and "refusers" in child psychotherapy: Reply to Garfield. *Journal of Consulting and Clinical Psychology, 57,* 170-171.

Weisz, J. R., Weiss, B., Morton, T., Granger, D., & Han, S. (1993). *Meta-analysis of psychotherapy outcome research with children and adolescents.* Unpublished manuscript, University of California at Los Angeles.

Wilson, G. T. (1985). Limitations of meta-analysis in the evaluation of the effects of psychological therapy. *Clinical Psychology Review, 5,* 35-47.

Witmer, H. L., & Keller, J. (1942). Outgrowing childhood problems: A study in the value of child guidance treatment. *Smith College Studies in Social Work, 13,* 74-90.