

# Effects of Psychotherapy With Children and Adolescents Revisited: A Meta-Analysis of Treatment Outcome Studies

John R. Weisz  
University of California, Los Angeles

Bahr Weiss  
Vanderbilt University

Susan S. Han and Douglas A. Granger  
University of California, Los Angeles

Todd Morton  
University of North Carolina, Chapel Hill

A meta-analysis of child and adolescent psychotherapy outcome research tested previous findings using a new sample of 150 outcome studies and weighted least squares methods. The overall mean effect of therapy was positive and highly significant. Effects were more positive for behavioral than for nonbehavioral treatments, and samples of adolescent girls showed better outcomes than other Age  $\times$  Gender groups. Paraprofessionals produced larger overall treatment effects than professional therapists or students, but professionals produced larger effects than paraprofessionals in treating overcontrolled problems (e.g., anxiety and depression). Results supported the specificity of treatment effects: Outcomes were stronger for the particular problems targeted in treatment than for problems not targeted. The findings shed new light on previous results and raise significant issues for future study.

Over the past decade, applications of the technique known as meta-analysis (see Cooper & Hedges, 1994; Mann, 1990; Smith, Glass, & Miller, 1980) have enriched our understanding of the impact of psychotherapy with children and adolescents (herein referred to collectively as “children”). At least three general meta-analyses encompassing diverse treatment methods and diverse child problems have indicated that the overall impact of child psychotherapy is positive, with effect sizes averaging not far below Cohen’s (1988) threshold of 0.80 for a “large” effect. Casey and Berman (1985) reported a mean effect size of 0.71 for a collection of treatment outcome studies with children 12 years of age and younger (studies published from 1952–1983). Weisz, Weiss, Alicke, and Klotz (1987) reported a mean effect size of 0.79 for a collection of studies with youth 4–18 years old (studies published from 1958–1984). Kazdin, Bass,

Ayers, and Rodgers (1990) assembled a collection of studies published between 1970 and 1988 focused on youth 4–18 years of age; mean effect sizes were 0.88 for treatment versus no-treatment comparisons and 0.77 for treatment versus active control comparisons. Combined, these three meta-analyses suggest that psychotherapy may have significant beneficial effects with children and adolescents (for a more detailed review of the meta-analytic procedures, findings, and implications, see Weisz & Weiss, 1993; Weisz, Weiss, & Donenberg, 1992).

In addition to generating overall estimates of therapy effectiveness, meta-analyses can provide evidence on factors that may be related to psychotherapy outcome. For example, meta-analytic evidence has fueled a lively debate over whether therapy outcomes differ as a function of the type of intervention used. Much of the adult therapy outcome literature has suggested that various forms of therapy work about equally well. One of the general conclusions reached by Smith et al. (1980) in their landmark meta-analysis was as follows:

Psychotherapy is beneficial, consistently so and in many ways. . . . Different types of psychotherapy (verbal or behavioral, psychodynamic, client-centered, or systematic desensitization) do not produce different types or degrees of benefit. . . . [It is] . . . clearly possible that all psychotherapies are equally effective, or nearly so; and that the lines drawn among psychotherapy schools are small distinctions, cherished by those who draw them, but all the same distinctions that make no important differences. (pp. 184, 186)

This conclusion has come to be called “the Dodo verdict” (i.e., “Everybody has won and all must have prizes”; see Parloff, 1984).

The picture may be different in the area of child psychotherapy, but the evidence is somewhat mixed thus far. Casey and Berman (1985) initially found that behavioral methods yielded significantly larger effect sizes than nonbehavioral methods;

---

John R. Weisz, Susan S. Han, and Douglas A. Granger, Department of Psychology, University of California, Los Angeles; Bahr Weiss, Department of Psychology and Human Development, Vanderbilt University; Todd Morton, Department of Psychology, University of North Carolina, Chapel Hill.

The work reported here was facilitated by National Institute of Mental Health (NIMH) Research Scientist Award K05 MH01161, NIMH Research Grant R01 MH49522, NIMH Grant 1-R18 SM50265, and Substance Abuse and Mental Health Services Administration Grant 1-HD5 SM50265-02.

We are grateful to Danika Kauneckis and Julie Mosk for their help with various phases of the project. We also thank the outcome study authors who provided us with information needed to calculate effect-size values for this meta-analysis.

Correspondence concerning this article should be addressed to John R. Weisz, Department of Psychology, Franz Hall, University of California, 405 Hilgard Avenue, Los Angeles, California 90024-1563. Electronic mail may be sent via Internet to weisz@psych.sscnet.ucla.edu.

however, when they excluded cases in which outcome measures were "very similar to activities occurring during treatment" (p. 391), the superiority of behavioral methods was reduced to a statistically marginal level ( $p = .06$ ; R. J. Casey, personal communication, July 30, 1992). Weisz et al. (1987) also found initially that, overall, behavioral treatments generated significantly larger effects than nonbehavioral treatments. Weisz et al. then eliminated only those outcome assessments for which outcome measures were unnecessarily similar to treatment activities. The rationale was that, with some treatments, the most valid test of efficacy may require an outcome measure that resembles the treatment activity; for example, when one treats a child's fear of dogs by graduated steps of approach, the most valid test of outcome is probably a behavioral test of the child's ability to approach dogs (i.e., an outcome measure that is necessarily similar to the treatment activities). For other treatments, by contrast, such resemblance is clearly unnecessary; for example, researchers who treat impulsivity through training on the Matching Familiar Figures Test (MFFT; see Kagan, 1965) and then use the MFFT to assess treatment outcome are using an unnecessarily similar outcome measure; it is not necessary to use the MFFT as an outcome measure because the goal of training is reduction of impulsivity, which can be measured in other, more ecologically valid ways than through readministration of the MFFT. After eliminating cases of unnecessary similarity, Weisz et al. (1987) found that behavioral therapies still generated significantly more positive effects than nonbehavioral approaches. Yet, meta-analytic findings on the relative effects of behavioral and nonbehavioral therapies remain a subject of controversy (e.g., see Weiss & Weisz, in press).

It is important that this issue and other controversial issues be addressed through new meta-analyses as new outcome studies are published. Given the inevitable variability of treatment effects across studies and the possibility of temporal shifts as methods evolve, it is necessary to assess the replicability of key findings across successive samples of treatment studies. As patterns are replicated across multiple meta-analytic samples, they may be accepted with increasing confidence. Of course, replication of findings across overlapping samples of studies may not be useful, because the overlap introduces a bias in favor of replication. Accordingly, in the present meta-analysis, we included only studies that had not been included previously in the two major comparative meta-analyses of child-adolescent psychotherapy research published thus far (i.e., Casey & Berman, 1985; Weisz et al., 1987).

Beyond the question of therapy method effects, it was important to test whether psychotherapy effects differ with the age of treated youngsters. Research on developmental differences in cognitive and social capacities (see Piaget, 1970; Rice, 1984) and on developmental differences in conformity to social norms and responses to adult authority (see Kendall, Lerner, & Craighead, 1984) suggests that, as children grow older, they may become less cooperative with adult therapists and less likely to adjust their behavior to societal norms. On the other hand, the cognitive changes (e.g., development of abstract thinking and hypotheticodeductive reasoning) that accompany maturation might also mean that, in contrast to children, adolescents better understand the purpose of therapy and are better suited to the verbal give-and-take that accompanies many forms of therapy

(for thoughtful discussions of diverse developmental issues related to child psychotherapy, see Holmbeck & Kendall, 1991; Shirk, 1988). However, the evidence on age and therapy outcome is mixed thus far. Casey and Berman (1985) found no relation between child age and study effect size, but Weisz et al. (1987) found a significant negative relation suggesting that older youth showed less positive changes in therapy than did their younger counterparts.

Several other previous findings need to be tested for robustness through further meta-analysis. Are girls more responsive to therapy than boys? The evidence is mixed thus far. Casey and Berman (1985) found that studies involving predominantly girls had better outcomes than studies involving mostly boys, but Weisz et al. (1987) found no significant relation between outcome and child gender. Do therapy effects differ for different treated problems? Here, too, the evidence is mixed. Casey and Berman found a lower mean effect size for social adjustment problems than for phobias, somatic problems, or self-control problems, but Weisz et al. found no reliable differences among such specific categories or between the broad categories of overcontrolled (e.g., phobias and social withdrawal) and undercontrolled (e.g., delinquency and noncompliance). Finally, meta-analytic comparisons can help one learn whether treatment outcomes are related to therapist experience. Casey and Berman found no relationship. However, although Weisz et al. also found no overall relationship, they did find two potentially informative interactions. Age and effect size were uncorrelated among children who saw professional therapists, but graduate students and paraprofessional therapists (e.g., trained teachers and parents) produced better outcomes with younger children than with older ones. In addition, professionals, students, and paraprofessionals did not differ in their success with undercontrolled children, but as amount of training and experience increased, so did effectiveness with overcontrolled children. Further evidence is needed to test the robustness of such effects.

A final substantive purpose of the current meta-analysis was to address the theoretically and practically important question of whether treatment effects with children are specific to the problems being addressed in treatment. Focusing on therapy with adults, Frank (1973) and others (see Brown, 1987) have argued that therapy has a variety of "nonspecific effects" (e.g., increasing the client's hopes and expectations of relief, producing cognitive and experiential learning, generating experiences of success, and enhancing feelings of being understood). Such speculation has led some outcome researchers (e.g., Bowers & Clum, 1988; Horvath, 1987) to explore the specificity of effects in psychotherapy (i.e., the extent to which an intervention's effects are specific to the theoretically targeted symptom domain vs. generalized across other symptom domains). Some (e.g., Horvath, 1988) have even taken the position that psychotherapy effects are "artifactual" to the extent that the therapy influences theoretically off-target symptoms. In the present study, we provided what appears to be the first test of treatment specificity in a broad-based meta-analysis of child outcome research. We tested whether treatment effects were larger for the specific target problems being addressed in therapy than for nontarget problems.

In addition to the substantive aims of the meta-analysis, we sought to address a little-noted but potentially critical statisti-

cal-analytic concern. Previous general meta-analysis of the child outcome literature have used the unweighted least squares (ULS) general linear model analytic approach (which subsumes regression and analysis of variance [ANOVA]) initially used by Smith et al. (1980).<sup>1</sup> For this approach to be precisely valid, several assumptions must be met, including the assumption of homogeneity of variance. This means that (a) the variance for the dependent variable for different subgroups (e.g., behavioral vs. nonbehavioral treatments) or within different ANOVA cells must be equivalent and (b) the variances of individual observations (i.e., in the case of meta-analysis, the individual effect sizes) must be equivalent across observations.

It may seem anomalous to refer to the variance of a single observation or, in the case of meta-analysis, the variance of a single effect size. However, in meta-analysis, the variance of an effect size (which refers to the reliability of that effect size) is directly computable and is, in part, a function of the sample size of the study on which the effect size is based. Because different studies often have quite different sample sizes, it is unlikely that the homogeneity assumption is satisfied in most meta-analyses (Hedges & Olkin, 1985).

To address this problem in the current analyses, we analyzed our data using the weighted least squares general linear models approach described by Hedges and Olkin (1985; see also Hedges, 1994), weighting effect sizes by the inverse of their variance. One drawback of our using this approach exclusively would be that meaningful comparison of the present findings with previous meta-analytic findings would be hampered; it would be impossible to determine whether differences between present and previous findings resulted from substantive differences in the relations among variables or from differences in analytic methods. Consequently, to facilitate comparison with previous findings, we report results for the current sample with both the WLS method, which we believe generates the most valid results, and the ULS general linear model method, which we (Weisz et al., 1987) and others (e.g., Casey & Berman, 1985) had used previously. To our knowledge, no such direct comparison has been made in previous meta-analyses.

### Method

As in our previous meta-analyses (Weiss & Weisz, 1990; Weisz et al., 1987), we defined *psychotherapy* as any intervention intended to alleviate psychological distress, reduce maladaptive behavior, or enhance adaptive behavior through counseling, structured or unstructured interaction, a training program, or a predetermined treatment plan. We excluded treatments involving drugs, interventions involving only reading (i.e., bibliotherapy), teaching or tutoring intended only to increase knowledge of a specific subject, interventions involving only relocation (e.g., moving children to a foster home), and exclusively preventive interventions intended to prevent problems in youngsters considered to be at risk. We included psychotherapy conducted by fully trained professionals, as well as psychotherapy conducted by therapists in training (e.g., clinical psychology and social work students and child psychiatry fellows) and by trained paraprofessionals (e.g., teachers and parents). Different schools of thought differ on the issue of whether extensive professional training is required for effective intervention. Here we treated this issue as an empirical question (see later discussion).

### Literature Search

We included only published psychotherapy outcome studies, relying on the journal review process as one step of quality control. We used

several approaches to identify relevant published studies (cf. Reed & Baxter, 1994; White, 1994). First, we conducted a computer search for the period January 1983 through April 1993, using the same 21 psychotherapy-related key terms used in Weisz et al. (1987) crossed with the same age-group constraints and outcome-assessment constraints used in that meta-analysis. Second, still focusing on the 1983–1993 time period, we also searched by hand, issue by issue, 30 journals that had produced articles in our 1987 meta-analysis. Third, and finally, we reviewed the list of studies included in previous meta-analyses conducted by Smith et al. (1980) and Kazdin et al. (1990) to identify any articles fitting our criteria that might have been missed in the other two steps of our search and that had not been included in the Casey and Berman (1985) or Weisz et al. (1987) meta-analysis. We did not exclude studies included in the Kazdin et al. (1990) survey and meta-analysis because we wanted to include tests of the impact of such factors as treatment method, treated problem, and therapist experience (tests that were not a focus of the Kazdin et al. report, which included only overall mean effect size values).

### Design and Reporting Requirements

To be included, a study had to include a comparison of a treated group with an untreated or attention control group. We excluded studies reporting only follow-up data (i.e., with no immediate posttreatment data). We excluded single-subject or within-subject designs. Such studies generate an unusual form of effect size (e.g., one based on intraparticipant variance, which is not comparable to conventional variance statistics) and, thus, do not appear to warrant equal weighting with studies that include independent treatment and control groups.

### Outcome Studies Generated by the Search

These several steps produced a pool of 150 studies<sup>2</sup> (marked with an asterisk in the References) involving 244 different treatment groups. The studies had been published between 1967 and 1993. Across the 150 studies, the mean age of the children ranged from 1.5 to 17.6 years ( $M = 10.50$ ,  $SD = 3.52$ ). Because the concept of “psychotherapy” with children 1.5 years old may seem implausible, we should note that studies focusing on very young children involved parent-training interventions focused on such early adjustment problems as frequent waking at night.

### Coding of the Studies

We coded studies for sample, therapy method, and design features, with parts of our coding system patterned after Casey and Berman (1985) and most of the system matching that used by Weisz et al. (1987); this overlap facilitated comparison with previous findings. Effect-size calculation and coding of the studies were carried out independently to avoid contamination. After training in the use of the coding system, four coders (i.e., four of the authors [two graduate students, one postdoctoral fellow, and one faculty member]) independently

<sup>1</sup> One exception to this generalization was a meta-analysis by Durlak, Fuhrman, and Lampman (1991), which focused on the effects of cognitive-behavioral interventions with children. Durlak et al. used a weighted least squares (WLS) group partitioning approach described by Hedges and Olkin (1985).

<sup>2</sup> Two of these studies were reported in a single article by Edleson and Rose (1982). Also, two Strayhorn and Weidman (1989, 1991) articles are reports of an original treatment study and its follow-up findings, respectively; thus, the two reports were coded as a single outcome study in our database.

Table 1  
Mean Effect Size and Significance Value (Versus 0) for Each Type of Therapy

Treatment type	n	Weighted least squares		Unweighted least squares	
		Effect size	p	Effect size	p
Behavioral	197	0.54	.0000	0.76	.0000
Operant	19	0.85 <sub>a</sub>	.0000	1.69 <sub>a</sub>	.0086
Physical reinforcement	5	1.67	.0000	2.43	.1119
Consultation in operant methods	2	1.59		2.17	
Social/verbal reinforcement	3	1.24	.0000	1.18	.3312
Self-reinforcement	2	1.29		3.57	
Combined physical/verbal reinf	5	0.85	.0000	1.38	.0616
Multiple operant methods	2	0.06		0.06	
Respondent	31	0.45 <sub>bc</sub>	.0000	0.70 <sub>b</sub>	.0005
Systematic desensitization	2	1.86		1.34	
Relaxation (no hierarchy)	23	0.41	.0000	0.72	.0022
Multiple respondent	3	0.42	.1554	0.54	.0959
Biofeedback	3	0.20	.3611	0.12	.3564
Modeling	12	0.40 <sub>bc</sub>	.0026	0.73 <sub>ab</sub>	.1018
Live nonpeer model	1	0.23		0.24	
Nonlive peer model	10	0.34	.0250	0.79	.1905
Nonlive nonpeer model	1	0.85		0.87	
Social skills	23	0.28 <sub>c</sub>	.0023	0.37 <sub>b</sub>	.0095
Cognitive/cognitive behavioral therapy	38	0.57 <sub>b</sub>	.0000	0.67 <sub>b</sub>	.0000
Parent training	36	0.49 <sub>bc</sub>	.0000	0.56 <sub>b</sub>	.0000
Multiple behavioral	35	0.61	.0000	0.86	.0015
Behavioral, not otherwise specified	3	0.33	.1507	0.38	.0486
Nonbehavioral	27	0.30	.0000	0.35	.0001
Client centered	6	0.11	.4218	0.15	.0884
Insight oriented	9	0.30	.0072	0.31	.0547
Discussion group	10	0.40	.0009	0.48	.0122
Nonbehavioral, not otherwise specified	2	0.46		0.38	
Mixed	20	0.63	.0000	0.55	.0002

Note. Tier 2 therapy types (e.g., operant and respondent) differed significantly for the behavioral categories ( $p < .01$  for both the weighted and unweighted least squares methods) but not for the nonbehavioral categories. Those Tier 2 behavioral categories that share a subscript did not differ significantly. For this table, as for results reported in the text, effect sizes were computed by collapsing up to the level of analysis (i.e., each study provided, at most, one effect size for each therapy category). However, in this table, sample sizes were computed without collapsing so as to provide information on the number of different treatment groups in our sample (i.e., the column labeled  $n$  shows number of treatment groups). Reinf = reinforcement.

coded 20% of the sample of studies. Mean interrater agreement ( $\kappa$ ) across pairs of coders is reported later for the various parts of the system.

**Therapy methods.** The three-tiered system shown in Table 1 was used in classifying therapy methods. Tier 1 included the broad categories behavioral and nonbehavioral (mean  $\kappa = .83$ ). Tier 2 included subcategories within each Tier 1 category (e.g., operant and client centered;  $\kappa = .75$ ). Tier 3 included finer grained descriptors applicable only to the behavioral methods (e.g., relaxation only vs. extinction only [within the Tier 2 respondent category];  $\kappa = .70$ ). We also coded each treatment as either group or individually administered ( $\kappa = .88$ ).

**Target problems.** Problems targeted by the treatment in the various studies were coded by means of the two-tiered system shown in Table 2. First, problems were grouped into either one of the two broadband categories most often identified in factor analyses of child behavior problems (e.g., see Achenbach & Edelbrock, 1978)—*overcontrolled* (e.g., social withdrawal) and *undercontrolled* (e.g., aggression)—or an *other category* ( $\kappa = .95$ ). Tier 2 included descriptive subcategories (e.g., delinquency, depression, and social relations) within each Tier 1 category ( $\kappa = .79$ ).

**Outcome measures.** Following Casey and Berman (1985), we coded whether outcome measures were similar to treatment activities ( $\kappa$

$= .79$ ). As noted earlier (and see detailed rationale in Weisz et al., 1987), we used an additional step of coding for outcome measures that were rated as similar to treatment activities; such measures were coded for whether the similarity was necessary (given the treatment goals) or unnecessary for a valid assessment ( $\kappa = .60$ ).

Following Casey and Berman (1985), we also coded outcome measures into *source* and *domain/type*. Source of outcome measure included such categories as *observers*, *therapists*, *parents*, and *participants' own performance* ( $\kappa = .88$ ). Domain included such content categories as *fear-anxiety*, *aggression*, and *social adjustment* ( $\kappa = .87$ ), grouped into broadband *overcontrolled* and *undercontrolled* categories. It may be useful to clarify the distinction between problem type and domain, both of which involved the same overcontrolled and undercontrolled broadband categories as well as the same Tier 2 codes. Problem type categories were used to classify the problems that were the target of the treatment intervention in a study, whereas domain categories were used to classify each of the individual outcome measures used in a study.

**Therapist training.** We classified therapists as to their level of professional training ( $\kappa = .87$ ). In our system, therapists were classified as *professionals* if they held a doctoral degree in medicine and had com-

Table 2  
*Mean Effect Size and Significance Value (Versus 0) for Each Type of Target Problem*

Target problem	<i>n</i>	Weighted least squares		Unweighted least squares	
		Effect size	<i>p</i>	Effect size	<i>p</i>
Undercontrolled	59	0.52	.00000	0.62	.00000
Delinquency	7	0.65 <sub>a</sub>	.00000	0.42	.00691
Noncompliance	4	0.39 <sub>ab</sub>	.04721	0.42	.05794
Self-control	18	0.44 <sub>a</sub>	.00002	0.87	.02053
Aggression	11	0.32 <sub>b</sub>	.00403	0.34	.00802
Multiple undercontrolled	19	0.49	.00000	0.67	.00374
Overcontrolled	40	0.52	.00000	0.69	.00000
Phobias/anxiety	16	0.60	.00000	0.57	.00106
Social withdrawal	4	0.66	.00243	0.71	.15998
Depression	6	0.64	.00011	0.67	.01076
Multiple overcontrolled	1	0.39		0.40	
Somatic (headaches)	13	0.37	.00009	0.86	.01034
Other	55	0.56	.00000	0.81	.00000
Academic				0.22	.02506
Low sociometric/peer acceptance	4	1.31 <sub>a</sub>	.00000	1.03	.08014
Social relations	11	0.47 <sub>b</sub>	.00000	1.12	.04571
Personality	5	0.56 <sub>b</sub>	.00071	0.53	.00343
Additional "other" problems	12	0.68	.00000	1.18	.01972
Multiple other	23	0.47	.00000	0.66	.00037

*Note.* Differences among Tier 2 problem type categories (e.g., self-control and aggression) were significant only with the weighted least squares method and only for two of the three broadband categories: undercontrolled ( $p < .05$ ) and other ( $p < .001$ ). Within each of those broadband categories, the Tier 2 categories that share a subscript did not differ significantly from one another. Note that academic problems were not included in weighted least squares analyses (see text for rationale) and that the Additional "other" problems and Multiple other categories were not included in two-way comparisons because they were too heterogeneous to be meaningful.

pleted psychiatric training or if they had completed a terminal doctorate or master's degree in psychology, education, or social work; they were classified as *graduate/professional students* if they were working toward professional credentials in psychiatry or toward advanced degrees in psychology, education, or social work; and they were classified as *paraprofessionals* (e.g., parents and teachers) if they lacked graduate training related to mental health but had been trained specifically to administer the therapy. For a group to be categorized at one of these three levels, at least two thirds of the therapists had to fit the level.

*Clinical versus analog samples.* We also coded studies for whether the samples were *clinical* (i.e., the youngsters were actually clinic referred or otherwise would have received treatment regardless of the study) or *analog* (i.e., the youngsters were recruited for the study and might not have received treatment had it not been for the study;  $\kappa = .89$ ).

## Results

### Overview of Data-Analytic Procedures

Before we report results, several data-analytic issues need to be considered.

*Confounding of independent variables.* A common problem in meta-analysis is that independent variables (e.g., therapy method and target problem) tend to be correlated (e.g., see Glass & Kliegl, 1983; Mintz, 1983). For example, because certain therapy methods tend to be used with certain target problems (e.g., behavioral methods with phobias), there is a natural confounding of therapy type and problem type. Here we used

an approach intended to address confounding while avoiding undue risk of either Type I or Type II error. Avoiding Type II error is especially important in meta-analyses given their potential heuristic, hypothesis-generating value.

In an initial wave of analysis, we conducted planned tests focused on five variables of primary interest: therapy type, target problem type, child age, child gender, and therapist training. For each variable, we first tested the simple main effect. Then we tested the robustness of the main effect using general linear models procedures to control statistically for the effects of each of the other four variables (see Appelbaum & Cramer, 1974); we covaried these variables individually in separate analyses, rather than simultaneously, to reduce the risk of capitalizing on chance. Finally, we tested each main effect for whether it was qualified by two-way or three-way interactions involving any of the other four variables (cf. Hedges, 1994).

*Effect-size computation.* As in our 1987 meta-analysis, we calculated effect sizes by dividing each study's posttherapy treatment group versus control group mean difference by the standard deviation of the control group. When means or standard deviations were not reported in the published article, we attempted to contact the author(s) to obtain the missing information. In those cases in which such efforts were unsuccessful, we used the procedures recommended by Smith et al. (1980) to derive effect-size values from inferential statistics such as *t* or *F* values; for reports of "nonsignificant" effects unaccompanied by any statistic, we followed the conservative procedure of esti-

rating the effect size at 0.00. In calculating effect-size values, some meta-analysts (e.g., Casey & Berman, 1985; Hedges, 1982) favor dividing by the pooled standard deviation of treatment and control groups, but others (e.g., Smith et al., 1980) believe that pooling is inappropriate because one effect of treatment may be to make variability greater in the treatment group than in the control group. Our previous analyses (see Weiss & Weisz, 1990) support the latter position. Nonetheless, for comparison purposes, we considered computing effect-size values twice in the present meta-analysis, once based on the pooled standard deviation and once based on the control group standard deviation. Before we could do this, it was necessary to test whether treatment and control group variances differed significantly. We found that, for the present sample of studies averaged within study, treatment and control group variances were reliably different (mean  $z = 0.23$ ,  $N = 79$ ,  $p < .05$ ) and, thus, that use of the pooled standard deviation was not appropriate. Accordingly, in calculating effect-size values for the present sample of studies, we divided by the unpooled control group standard deviation. Throughout our calculations, we consistently collapsed across outcome measures and treatment groups up to the level of analysis. For example, we collapsed across treatment groups except when we tested the effects of therapy type.

As noted earlier, we analyzed our data using two different approaches to permit comparisons between previous and present results unconfounded with analytic differences. The primary analyses involved the WLS approach, with each effect size weighted by the inverse of its variance (Hedges, 1994; Hedges & Olkin, 1985); this adjusted for heterogeneity of variance across individual observations. In the WLS analyses, we did not include academic outcomes (e.g., school grades), and we dropped outliers (i.e., effect sizes lying beyond the first gap of at least one standard deviation between adjacent effect-size values in a positive or negative direction; Bollen, 1989).<sup>3</sup> We dropped academic outcomes because so many factors (e.g., intelligence) other than psychopathology could be responsible for poor academic performance that it seemed inappropriate to base tests of psychotherapy efficacy on such outcomes. We dropped outliers so that our results would fairly represent the large majority of the data. Our secondary analyses, included for comparison purposes only, used the ULS approach (with academic outcomes and outliers included) used in previous meta-analyses conducted by Casey and Berman (1985), Kazdin et al. (1990), Smith et al. (1980), and Weisz et al. (1987).

### *Overall Mean Effect Size: WLS and ULS Findings*

When we used the WLS method of analysis, with one effect size per study, the mean effect size across the 150 studies was 0.54 (significantly different from 0),  $\chi^2(1, N = 150) = 404.65$ ,  $p < .000001$ . When we used the ULS method, again with one effect size per study, the mean effect size was 0.71 (significantly different from 0),  $t(149) = 9.49$ ,  $p < .000001$ . The latter mean was comparable to the ULS-derived means reported in earlier child meta-analyses conducted by Casey and Berman (1985;  $M$  effect size = 0.71), Weisz et al. (1987;  $M$  effect size = 0.79), and Kazdin et al. (1990;  $M$  effect size = 0.88 for treatment vs. nontreatment comparisons, and  $M$  effect size = 0.77 for treatment vs. active control groups). As the WLS and ULS figures

illustrate, one effect of the WLS adjustment is that it often produces lower mean effect-size values when one averages across multiple studies; larger effect-size values tend to have smaller weights because they usually involve smaller sample sizes and larger variances than do smaller effect-size values.

### *Therapy Type: Behavioral Versus Nonbehavioral Interventions*

*Primary analysis with WLS.* The mean effect size with the WLS method was higher for behavioral therapies than for nonbehavioral therapies (effect-size means: 0.54 vs. 0.30),  $\chi^2(1, N = 149) = 9.96$ ,  $p < .005$ . The difference remained significant when we controlled for age, gender, and therapist training and was marginally significant ( $p < .10$ ) when we controlled for problem type. This main effect of therapy type supports the hypothesis that behavioral treatments are more effective than nonbehavioral treatments. However, it is possible that the effects of behavioral treatments are less enduring than those of nonbehavioral treatments. If this were the case, the apparent superiority of behavioral treatments would be significantly qualified. We tested this possibility by comparing the difference between posttreatment and follow-up effect sizes for behavioral versus nonbehavioral treatments. The test was nonsignificant, indicating that behavioral and nonbehavioral methods did not differ in the durability of their effects.

All two-way interactions between therapy type and the other four variables were nonsignificant. However, the Therapy Type  $\times$  Child Age  $\times$  Therapist Training interaction was significant,  $\chi^2(1, N = 99) = 9.49$ ,  $p < .01$ . Of the component two-way interactions, only two were significant. First, among paraprofessional therapists only, therapy type and child age interacted ( $p < .005$ ), indicating that paraprofessionals treating children (but not adolescents) produced a larger mean effect size with behavioral than with nonbehavioral therapies (means: 1.2 vs. 0.46,  $p < .005$ ) and that paraprofessionals using behavioral (but not nonbehavioral) methods produced larger effect sizes with children than with adolescents (means: 1.20 vs. 0.03,  $p < .0001$ ). The second component two-way interaction was Professional Training  $\times$  Child Age for studies involving behavioral therapies,  $\chi^2(2, N = 81) = 51.37$ ,  $p < .0001$ . Simple effects tests revealed that, for adolescents treated with behavioral methods, student therapists achieved larger effect sizes than did paraprofessionals ( $p < .05$ ); in treatment of children, none of the pairwise differences between levels of therapist experience were significant. Also, paraprofessionals using behavioral methods achieved larger effect sizes for children than for adolescents (means: 1.2 vs. 0.03,  $p < .0001$ ), but the reverse was true when behavioral methods were used by professionals (means: 0.91 vs. 0.55,  $p < .05$ ) and student therapists (means: 0.68 vs. 0.33,  $p < .01$ ). Across the 12 means encompassed by this triple interaction, the largest mean effect size was achieved by paraprofessionals using behavioral methods with children (1.20); the

<sup>3</sup> There were 11 outlier effect-size values ranging from 7.89 to 28.00. All of the outliers came from three analog sample studies using behavioral treatments. Beyond this, there was no particular pattern to the outliers that we could discern.

smallest mean effect size involved paraprofessionals using behavioral methods with adolescents (0.03).

*Secondary analysis with ULS.* The mean effect size was also higher for behavioral therapies than for nonbehavioral therapies when we used the traditional ULS method. With the Glass method, the difference in means (0.76 vs. 0.35) was significant,  $F(1, 147) = 4.09, p < .05$ . The difference remained significant when we adjusted for age and therapist training and was marginally significant ( $p < .10$ ) when we controlled for gender and problem type. Interactions between therapy type and each of these other four variables were nonsignificant. Mean effect sizes for the various types of therapy are shown in Table 1.

*Adjusting for similarity of treatment activities and outcome measures: WLS.* As expected, outcome measures coded as similar to treatment activities (described earlier) generated larger effect-size values than measures coded as dissimilar (means: 0.67 vs. 0.48),  $\chi^2(1, N = 189) = 10.56, p < .001$ . When we excluded all similar outcome measures, the therapy type effect remained significant (means: 0.47 vs. 0.25),  $\chi^2(1, N = 133) = 8.18, p < .005$ . When we excluded only outcome measures coded as unnecessarily similar, behavioral methods showed more pronounced superiority to nonbehavioral methods (means: 0.52 vs. 0.25),  $\chi^2(1, N = 148) = 13.12, p < .0005$ .

*Adjusting for similarity with ULS.* Using the ULS method, we again found that outcome measures similar to treatment activities showed larger effect sizes than dissimilar outcome measures (means: 0.98 vs. 0.56),  $F(1, 187) = 8.18, p < .005$ . When we excluded all outcome measures that were coded as similar, the mean effect size for behavioral treatments (0.59) was marginally higher ( $p < .10$ ) than that for nonbehavioral treatments (0.30). However, as in the Weisz et al. (1987) analysis, when we excluded only outcome measures that were coded as unnecessarily similar, behavioral treatments showed a significantly higher mean effect size than nonbehavioral approaches (means: 0.73 vs. 0.30),  $F(1, 146) = 3.86, p < .05$ .

### *Problem Type: Overcontrolled and Undercontrolled*

*Primary analysis with WLS.* Table 2 shows effect-size means for the various target problems. Using the WLS method, we compared means for the overcontrolled and undercontrolled categories (we dropped the "other" category because it encompassed a heterogeneous group of problems with no clear theoretical referent). We found no significant problem type main effect. This finding suggests that treatment may be about equally effective with overcontrolled and undercontrolled problems. However, it is possible that, although the two types of problems are similarly responsive to treatment in the short term, treatment effects are more enduring for one of the problem domains than for the other. To test this possibility, we compared the difference between posttreatment effect size and follow-up effect size for overcontrolled versus undercontrolled problems; the test was nonsignificant, indicating no difference in the stability of treatment effects for the two types of problems.

The main effect of problem type remained nonsignificant when we controlled for therapy type, age, and therapist training. However, when we controlled for gender of treated children, the problem type effect was marginally significant ( $p < .10$ ), with a marginally larger mean effect size for undercontrolled (0.58) than overcontrolled (0.44) problems. Interactions between

problem type and therapy type, age, and gender were all nonsignificant. However, the Problem Type  $\times$  Therapist Training interaction was significant,  $\chi^2(2, N = 59) = 22.24, p < .00001$ . The effect of problem type was marginally significant for students ( $p < .10$ ) and significant for paraprofessionals,  $\chi^2(1, N = 9) = 13.34, p < .001$ , and professionals,  $\chi^2(1, N = 25) = 5.87, p < .05$ . Paraprofessionals were more effective with undercontrolled than overcontrolled children (effect-size means: 0.72 and 0.14), whereas the reverse was true of professionals (effect-size means: 0.49 and 0.86). Analyzing the interaction from another perspective, we found that the effect of therapist training was significant for both internalizing ( $p < .001$ ) and externalizing ( $p < .005$ ) problems, but the pattern differed for the two types of problems. In treating undercontrolled problems, paraprofessionals ( $M$  effect size = 0.72) produced larger effects than professionals ( $p < .05$ ;  $M$  effect size = 0.49) and students ( $p < .001$ ;  $M$  effect size = 0.29); professionals and students did not differ reliably. By contrast, in treating overcontrolled problems, professionals ( $M$  effect size = 0.86) generated larger effects than paraprofessionals ( $p < .001$ ;  $M$  effect size = 0.14) and marginally larger effects than students ( $p < .10$ ;  $M$  effect size = 0.56), and students were significantly superior to paraprofessionals ( $p < .05$ ).

The Problem Type  $\times$  Child Age  $\times$  Therapist Training interaction was also significant,  $\chi^2(2, N = 59) = 8.24, p < .05$ . Three of the component two-way interactions were significant. First, for adolescents only, problem type interacted with therapist training,  $\chi^2(2, N = 22) = 18.36, p < .0001$ . Simple effects tests showed that (a) when adolescents were treated by paraprofessionals (but not other therapist groups), the effect size was higher for undercontrolled ( $M = 0.75$ ) than for overcontrolled ( $M = 0.01$ ) problems ( $p < .0001$ ), and (b) when adolescents were treated for overcontrolled (but not undercontrolled) problems, professionals achieved a higher effect size than students, who in turn had a higher effect size than paraprofessionals (means: 1.33 vs. 0.78 vs. 0.01,  $p < .0001$ ). The second component two-way interaction was Therapist Training  $\times$  Child Age for youth with overcontrolled problems,  $\chi^2(2, N = 25) = 11.33, p < .005$ . Simple effects tests showed that, for overcontrolled youth treated by professionals and students (but not by paraprofessionals), the effect size was higher for adolescents than for children (for professionals, means of 1.33 and 0.66,  $p < .05$ ; for students, means of 0.78 and 0.10,  $p < .005$ ). Also, as noted earlier, when adolescents were treated for overcontrolled problems, professionals generated larger effects than students, who in turn generated larger effects than paraprofessionals. The third component two-way interaction was Problem Type  $\times$  Child Age for studies using paraprofessional therapists; as noted earlier, paraprofessionals treating adolescents achieved larger effects with undercontrolled than overcontrolled target problems. Across the 12 effect-size means encompassed by this triple interaction, the smallest (0.01) was produced by paraprofessionals treating overcontrolled adolescents, whereas the largest (1.33) was produced by professionals treating the same group.

*Secondary analysis with ULS.* The ULS method produced rather different results. The main effect of problem type was nonsignificant, and it remained so when we controlled for therapy type, age, gender, and therapist training. When we tested for interactions between problem type and these four variables, we

found only a marginal Problem Type  $\times$  Gender interaction ( $p < .10$ ).

### Child Age

*Primary analysis with WLS.* When we examined age differences using the WLS method, we found a larger mean effect for studies treating adolescents (12 years of age and older;  $N = 37$  studies,  $M$  effect size = 0.65) than for studies treating children (11 years of age and younger;  $N = 110$ ,  $M$  effect size = 0.48),  $\chi^2(1, N = 147) = 9.08, p < .01$ . The age main effect remained significant when we controlled for problem type and gender; however, the effect became marginal ( $p < .10$ ) when we controlled for therapy type and nonsignificant when we controlled for therapist training. The Age  $\times$  Therapy Type and Age  $\times$  Problem Type interactions were nonsignificant, but age showed significant interactions with gender,  $\chi^2(1, N = 121) = 15.20, p < .0001$ , and therapist training,  $\chi^2(2, N = 99) = 17.96, p < .0001$ .

Analyzing the Age  $\times$  Gender interaction, we found that the effect of age was nonsignificant for male majority samples but significant for female samples ( $p < .00001$ ; effect-size means of 0.86 for adolescents and 0.43 for children). Viewing this interaction from the other direction, we found that the gender effect was significant among adolescents ( $p < .00001$ ; female  $M = 0.86$ , male  $M = 0.37$ ) but not among children. The interaction is shown in Figure 1.

Analyzing the Age  $\times$  Therapist Training interaction, we found that treatment effects were stronger for adolescents than children when the treatment was provided by professionals ( $p < .01$ ; effect-size means of 0.87 and 0.47) and by students ( $p < .01$ ; effect-size means of 0.66 and 0.33); the reverse was true, however, when treatment was provided by paraprofessionals ( $p$

$< .05$ ; effect-size means of 0.62 and 0.92). Viewing this interaction from the other direction, we found that therapist training had a significant effect among children ( $p < .00001$ ) but not among adolescents. Among children, the effect size was higher for paraprofessionals than for professionals ( $p < .001$ ) and student therapists ( $p < .00001$ ), and the last two groups were not significantly different.

*Secondary analysis with ULS.* ULS analyses showed a nonsignificant age main effect that remained nonsignificant when we controlled for therapy type, problem type, gender, and therapist training. The interactions between age and therapy type, problem type, and gender were also nonsignificant; however, the Age  $\times$  Therapist Training interaction was significant,  $F(2, 94) = 5.10, p < .01$ . Analyzing this interaction, we found that the effect of age was nonsignificant for students and paraprofessionals but marginally significant for professionals,  $F(1, 45) = 3.93, p < .10$ , who had better outcomes with adolescents ( $M$  effect size = 1.05) than with children ( $M$  effect size = 0.59). Viewing the interaction from the alternate direction, we found that there was a significant therapist training effect with children (effect-size means of 1.61 for paraprofessionals, 0.61 for professionals, and 0.42 for students),  $F(2, 70) = 6.86, p < .005$ , but no significant effect with adolescents.

### Child Gender

*Primary analysis with WLS.* To examine the relation between effect size and child gender, we categorized studies into those in which most participants were male ( $n = 78$ ) and those in which most participants were female ( $n = 45$ ). In our primary analysis, using the WLS method, we found a significant gender main effect,  $\chi^2(1, N = 122) = 20.04, p < .00001$ , reflecting the fact that the mean effect size was higher for female

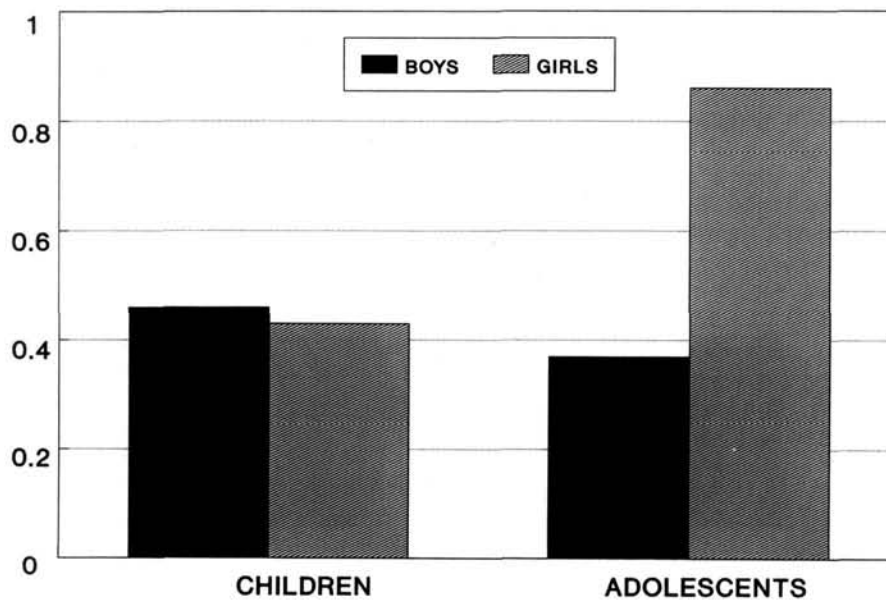


Figure 1. Mean effect size for samples of predominantly male and female children (11 years of age and younger) and adolescents (12 years of age and older).

(0.71) than male (0.43) samples. This effect was reduced to a marginal level when we controlled for therapy type ( $p < .10$ ) but remained significant when we controlled for problem type, age, and therapist training. Gender did not show significant interactions with therapy type or problem type, but there were two significant interactions involving gender. One was the Gender  $\times$  Age interaction discussed earlier (see Figure 1). The other was a Gender  $\times$  Therapist Training interaction,  $\chi^2(2, N = 77) = 9.36, p < .01$ . This reflected, in part, the fact that the gender effect was significant among youngsters treated by professionals ( $p < .001$ ; female  $M = 0.93$ , male  $M = 0.41$ ) and paraprofessionals ( $p < .0001$ ; female  $M = 0.99$ , male  $M = 0.49$ ) but not among those treated by students. Viewing this interaction from the other direction, we found that therapist training had a significant effect in female samples ( $p < .0001$ ) but not in male samples. In the female samples, student therapists had a lower mean effect size (0.48) than professionals (0.93,  $p < .01$ ) and paraprofessionals (0.99,  $p < .0001$ ), but the last two groups did not differ significantly.

*Secondary analysis with ULS.* Our secondary, ULS analysis revealed a marginal gender main effect ( $p < .10$ ) reflecting a trend toward larger effects for female samples ( $M$  effect size = 0.80) than for male samples ( $M$  effect size = 0.54). This effect remained marginally significant when we controlled for therapy type, became nonsignificant when we controlled for problem type and age, and became significant ( $p < .05$ ) when we controlled for therapist training. Interactions between gender and therapy type and age were nonsignificant, and interactions between gender and problem type and therapist training were marginally significant ( $p < .10$ ).

### Level of Therapist Training

*Primary analysis with WLS.* In our primary, WLS analysis, we found a significant main effect of therapist training,  $\chi^2(2, N = 100) = 12.51, p < .005$ , with paraprofessionals ( $M$  effect size = 0.71) generating significantly larger effects than professionals ( $p < .05$ ;  $M$  effect size = 0.55) and students ( $p < .001$ ;  $M$  effect size = 0.43) and the last two groups not differing significantly. This therapist training main effect remained significant when we controlled for age and gender but was reduced to a marginal level ( $p < .10$ ) when we controlled for therapist type and problem type. The Therapist Training  $\times$  Therapy Type interaction was nonsignificant, but therapist training interacted significantly (as described earlier) with problem type, age, and gender.

We tested whether the main effect involving therapist training might have been influenced by a tendency of less fully trained therapists to be assigned to analog–nonclinic cases (see other analog vs. clinic analyses detailed later). A 2 (analog vs. clinic cases)  $\times$  3 (therapist training) chi-square analysis revealed a marginal trend in this direction,  $\chi^2(2, N = 98) = 5.52, p < .10$ , but this trend did not account for the main effects of therapist training reported earlier. The main effect remained significant when we controlled for the analog versus clinic variable using both the WLS method ( $p < .005$ ) and the ULS method ( $p < .05$ ). And the interaction between therapist training and analog–clinic was nonsignificant in both WLS and ULS analyses. Findings with regard to therapist training and the other four variables of primary interest are summarized in Table 3.

*Secondary analysis with ULS.* In secondary analyses with the ULS method, the main effect of therapist training (professional vs. paraprofessional vs. student therapist) was again significant,  $F(2, 97) = 3.26, p < .05$ ; paraprofessionals generated larger effect sizes ( $M = 1.25$ ) than professionals ( $M = 0.70, p < .05$ ) and students ( $M = 0.57, p < .05$ ), whereas the latter two groups did not differ significantly. This main effect of therapist training remained significant when we controlled for therapy type and age, but it became nonsignificant when we controlled for gender and problem type. Interactions of therapist training with therapy type and with problem type were nonsignificant, but therapist training did show a significant interaction with age (described earlier) and a marginal interaction with gender ( $p < .10$ ).

### Other Potential Correlates of Effect Size

In a secondary wave of analyses, we examined the relation between effect size and other variables that we believed might help to explain the results reported earlier.

*Source and domain of outcome measure.* We tested effects of the source of outcome measures (parents,  $n = 43$ ; teachers,  $n = 58$ ; observers,  $n = 43$ ; peers,  $n = 13$ ; participant performance,  $n = 53$ ; and self-report,  $n = 81$ ) and the domain of outcome measures (grouped into overcontrolled and undercontrolled). We also tested the interaction between source and domain, thinking that different sources might be differentially sensitive to child behavior within the overcontrolled and undercontrolled categories. Neither main effect was significant in either our primary (WLS) or secondary (ULS) analysis. The interaction of the two, however, was significant in our primary, WLS analysis (although not in the secondary, ULS analysis),  $\chi^2(5, N = 185) = 13.56, p < .05$ . The effect of domain was nonsignificant for parents and self-report and marginally significant ( $p < .10$ ) for each of the other sources. The effect of source was significant for both the undercontrolled ( $p < .005$ ) and overcontrolled ( $p < .005$ ) domains, but the pattern across the various sources was quite different for the two problem domains, as shown in Figure 2. In essence, positive treatment effects on overcontrolled problems were most likely to be obtained through reports by peers and by treated children themselves, whereas positive effects on undercontrolled problems were more likely to be obtained through reports by teachers and trained observers and through direct behavioral assessment.

*Analog versus clinic cases.* We compared effect sizes for studies involving clinic children ( $n = 41$ ) and for studies involving analog or recruited cases ( $n = 104$ ; 5 were uncodable). In our primary, WLS analysis (but not in our secondary, ULS analysis), the groups were found to differ reliably,  $\chi^2(1, N = 145) = 6.68, p < 0.01$ ; analog samples ( $M = 0.59$ ) generated larger effect sizes than clinic samples ( $M = 0.43$ ).

*Group-administered versus individual treatment.* We compared group interventions ( $n = 92$ ) and individual interventions ( $n = 41$ ; we dropped combined and milieu treatments). There was a significant effect-size mean difference in our primary, WLS analysis (but not in our secondary, ULS analysis), showing a higher mean effect size for individually administered

Table 3  
Summary of Meta-Analysis Findings for the Five Variables of Primary Interest

Variable	Main effect		Significant after control		Significant interaction with	
	WLS	ULS	WLS	ULS	WLS	ULS
Therapy type	Behavioral > nonbehavioral***	Behavioral > nonbehavioral**	Problem type Child age Child gender Therapist training	Problem type Child age Child gender <sup>a</sup> Therapist training		
Problem type			Child gender <sup>a</sup> Therapy type <sup>a</sup> Problem type Child gender		Therapist training***** Child gender***** Therapist training***	Child gender* Therapist training**
Child age	Adolescents > children***		Therapy type <sup>a</sup> Problem type Child age Therapist training	Therapy type <sup>a</sup> Therapist training	Child age***** Therapist training***	Problem type* Therapist training*
Child gender	Female > male*****	Female > male*	Therapy type <sup>a</sup> Problem type Child age Therapist training	Therapy type Problem type Child age	Problem type***** Child age**** Child gender***	Child age*** Child gender**
Therapist training	Paraprofessional > professional > student***	Paraprofessional > professional = student**	Therapy type <sup>a</sup> Problem type <sup>a</sup> Child age Child gender			

Note. Listings in the column labeled Significant after control indicate that the main effect shown at left remained significant when the variable listed in the column was controlled. WLS = weighted least squares; ULS = unweighted least squares.

<sup>a</sup> Main effect became marginal ( $p < .10$ ) when variable was controlled.

\*  $p < .10$ . \*\*  $p < .05$ . \*\*\*  $p < .01$ . \*\*\*\*  $p < .001$ . \*\*\*\*\*  $p < .0001$ .

treatments (0.63) than for group treatments (0.50),  $\chi^2(1, N = 133) = 4.36, p < .05$ .

*Posttreatment versus follow-up.* Across outcome measures that involved both immediate posttreatment and longer term follow-up assessments, we compared mean effect sizes for post-treatment and follow-up. Across the 50 studies that included follow-up assessment, the mean lag between posttreatment and

follow-up was 28.4 weeks. We found no significant difference in either our primary (WLS) or secondary (ULS) analysis.

#### Year-of-Publication Effects

Finally, we conducted two analyses that had not been done in previous meta-analyses and thus did not require a ULS com-

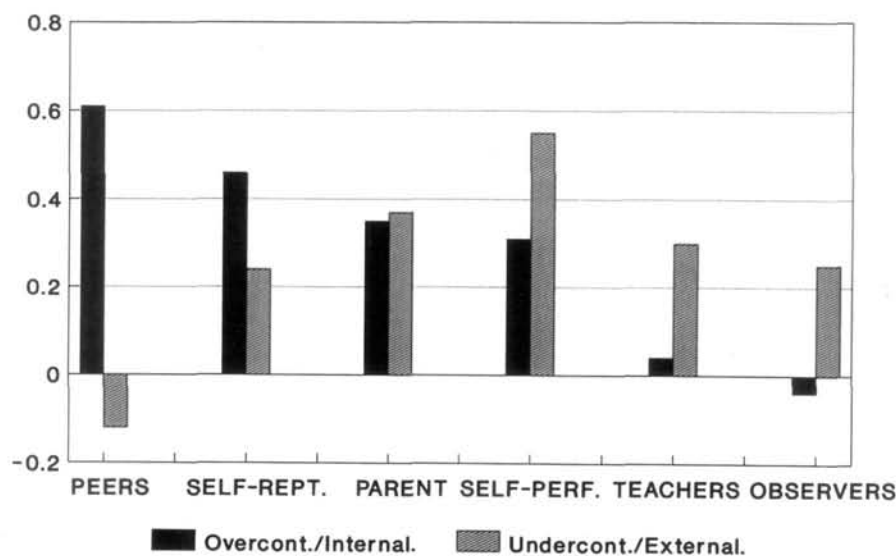


Figure 2. Mean effect size as a function of source and domain of outcome measure. REPT. = report; PERF. = performance; Overcont. = overcontrolled; Internal. = internalized; Undercont. = undercontrolled; External. = externalized.

parison; thus, for these last two analyses, we report only WLS findings. In the first such analysis, we assessed whether year of publication was related to effect size, to test whether such a relationship might explain differences between the present findings and those of Weisz et al. (1987). Accordingly, we divided the studies into pre-1985 ( $n = 60$ ) and 1986–1993 ( $n = 91$ ) subsets (1985 was the cutoff date for the sample of studies included in Weisz et al., 1987). In our primary analysis, using the WLS method, the main effect for year was nonsignificant; also nonsignificant were the interactions between year and therapy type, problem type, and therapist training, indicating that the relation between those variables and therapy outcome had not changed appreciably over time. However, there were significant interactions with age,  $\chi^2(1, N = 147) = 5.34, p < .05$ , and gender,  $\chi^2(1, N = 122) = 7.21, p < .01$ . The effect of year was nonsignificant for children but significant for adolescents ( $p < .05$ ), for whom more recent studies had a larger mean effect size (0.69 vs. 0.46). Also, the effect of year was nonsignificant for boys but significant for girls ( $p < .05$ ), for whom recent studies had a larger mean effect size (0.75 vs. 0.41).

### *Specificity of Treatment Effects*

Finally, as noted in the introduction, we addressed the important question of whether treatment has broad, general effects or whether treatment produces effects that are specific to the target problems being addressed. We assessed whether the mean effect size was larger for outcome measures that were within the same broad domain (i.e., overcontrolled vs. undercontrolled; see Table 2) as the target problem being treated than for outcome measures in a different broad domain than the target problem; it was. Matches (i.e., instances in which target problem and outcome measure fell into the same broad domain) were, in fact, associated with a significantly larger mean effect size (0.52) than were outcome measures that fell into different broad domains than the target problem (0.22),  $\chi^2(1, N = 101) = 8.40, p < .005$ .

To determine the level of specificity at which this effect occurred, we assessed whether matching precisely within the broadband factors (e.g., matching outcome measures and target problems at such specific levels as anxiety, depression, and social withdrawal) was associated with larger effects than simply matching on broadband factors (e.g., overcontrolled). That is, we compared the mean effect size for outcome measures that matched target problems precisely (e.g., an outcome measure of anxiety when anxiety was the focus of treatment) and the mean effect size for equally specific outcome measures that matched targeted problems only at the broadband level (e.g., an outcome measure of depression when anxiety was the focus of treatment; both anxiety and depression are overcontrolled problems, but only one precisely matches the focus of treatment). In this analysis, we excluded all outcome measures that did not match target problems at the broadband level. Also, for this analysis, unlike that discussed in the previous paragraph, we included measures within the broadband “other” category because it was not heterogeneous at the narrow band level. The comparison showed that outcome measures that precisely matched target problems yielded a markedly larger mean effect size (0.60) than did equally specific outcome measures

that did not match target problems (0.30),  $\chi^2(1, N = 180) = 33.43, p < .000001$ . This suggests that the effects of treatment across the studies we examined were specific to the domains targeted in treatment.

### *Recomputing ULS Analyses*

For all of the ULS calculations reported earlier, we used the procedures of previous meta-analyses so as to facilitate comparisons with previous findings. Accordingly, we included academic outcomes, and we did not exclude outliers on any of the outcome measures used. Although this procedure was necessary for comparison with previous findings, it complicates comparison of our current ULS findings with our current WLS findings because, for our WLS analyses, we excluded all academic outcome measures and all outliers. To address this limitation, we reran all ULS analyses, this time excluding all outliers (i.e., effect sizes lying beyond the first gap of at least one standard deviation between adjacent effect-size values; Bollen, 1989) and all academic outcomes. This change in procedure had very little effect on our ULS results. The overall mean effect size changed from 0.71 to 0.69, and a few significance levels changed: (a) The main effect of problem type remained nonsignificant in nearly all ULS analyses but became significant when we controlled for therapist training ( $p < .05$ ;  $M$  undercontrolled effect size = 0.42,  $M$  overcontrolled effect size = 0.89); (b) the main effect of gender changed from marginal to significant ( $p < .05$ ;  $M$  male effect size = 0.53,  $M$  female effect size = 0.84) and remained significant with therapy type controlled (after having previously been marginally significant); and (c) the main effect of therapist training changed from significant ( $p < .05$ ) to nonsignificant and remained so with therapy type and then age controlled. With these few exceptions, all previously nonsignificant effects remained nonsignificant, and all previously significant effects remained significant and at the same level. The strong similarity indicates that the numerous differences reported earlier between ULS and WLS findings cannot be attributed to our having dropped outliers and academic outcomes in the ULS analyses.

### Discussion

The current findings, combined with results of previous meta-analyses, provide convergent evidence on several key issues in child and adolescent psychotherapy. First, and most generally, the present findings reinforce previous evidence that psychotherapy with young people produces positive effects of respectable magnitude. On the other hand, the mean effect-size values generated by our primary analyses (using WLS methods) suggest that previous estimates of the magnitude of treatment effects have probably been inflated relative to what we believe to be the fairest estimate of effect magnitude in the population. Previous meta-analytic estimates, like our present ULS analysis, have generated mean effect-size values in the range just below Cohen's (1988) standard of 0.80 for a large effect; our current best estimate, based on the WLS method, was 0.54, closer to Cohen's standard of 0.50 for a medium effect. The tendency of weighted analyses to produce smaller effects than unweighted approaches may not be confined to

child psychotherapy research but may instead reflect a general tendency for treatment effects to be inversely related to sample size. Robinson, Berman, and Neimeyer (1990) reported this effect in their review of depression treatment research, suggesting that it may result from publication bias in favor of statistically significant treatment effects. That is, when sample sizes are small, only large effects will be statistically significant and thus likely to be published.

The adult outcome literature has provided some support for "the Dodo verdict," the notion that different approaches to therapy are about equally effective. The present findings do not support such a conclusion with respect to treatment of children. We found that behavioral methods were associated with more substantial therapy effects than nonbehavioral methods, and this pattern held up when we controlled for outcome measures that were unnecessarily similar to treatment activities. The pattern also held up when we controlled for treated problem, therapist training, and child age and gender, and the main effect of behavioral versus nonbehavioral methods was not qualified by interactions with any of those four variables. Superior effects of behavioral over nonbehavioral interventions were also reported in earlier child meta-analyses conducted by Casey and Berman (1985; but see qualifications in the introduction and in Weisz & Weiss, 1993) and Weisz et al. (1987). Because none of the 150 studies in the present meta-analysis was included in either the Casey and Berman (1985) or Weisz et al. (1987) analysis, the present findings must be seen as rather strong independent evidence of the replicability of this "non-Dodo verdict." On the other hand, only about 10% of the treatment groups in our sample involved nonbehavioral interventions; in the future, it will be important for researchers to expand the base of evidence on the effects of the nonbehavioral interventions that are widely used in clinical practice but rarely evaluated for their efficacy in controlled studies.

In addition, we must stress that even highly significant behavioral-nonbehavioral differences in outcome may ultimately prove to have artifactual explanations. Previous analyses with predominantly adult samples have shown a significant relation between investigator allegiance—or expectancies as to which therapy method will be more successful—and outcome when different therapeutic methods are being compared (see Berman, Miller, & Massman, 1985; Robinson et al., 1990; Smith et al., 1980). Such findings do not clearly answer the question of whether investigator expectancy causes—or results from—findings of behavioral-nonbehavioral comparisons or whether both processes are operative. Certainly, investigators' expectancies regarding the relative effectiveness of treatments do not develop in a vacuum but are likely to be based on their past experience and on the results of previous studies, their own as well as others. This fact makes it difficult to "control for" investigator expectancies statistically (e.g., by testing for behavioral vs. nonbehavioral outcome differences with investigator expectancies covaried).

To illustrate the problem, we offer an analogy. Suppose that surgeons over the years have used either Surgical Method A or Surgical Method B to address a particular type of heart ailment. Over these years, increasing evidence has indicated that Method A produces superior survival rates. Suppose, further, that a reviewer conducts a statistical analysis "controlling for" the ten-

dency of surgeons in comparative studies to believe that A is more effective than B, and the reviewer finds that, with that belief "controlled," there is no remaining reliable difference between Methods A and B. Would this really mean that A is no more effective than B? No. It would mean only that the belief that A is more effective than B corresponds so closely to the evidence that A is more effective than B that controlling for the belief also controls for the evidence. In such a case, applying statistical control could be a disservice to the field.

In our view, the potential role of investigator expectancies or allegiance has been highlighted sufficiently in previous correlational research that it warrants attention in future research. But the only fair way we know to study the role of such expectancies is experimentally, by judicious assignment of therapists to conditions. For example, outcome researchers who wish to compare behavioral and nonbehavioral methods might use factorial designs in which therapy methods are crossed with therapist orientation (behavioral vs. nonbehavioral). We suspect that the results of such research would contribute more to understanding than a proliferation of difficult-to-interpret correlational analyses.

In contrast to our therapy type findings, we found no evidence that therapists had different levels of success with overcontrolled problems than with undercontrolled problems. The same finding was obtained in the only other meta-analysis to address this question (Weisz et al., 1987). Evidence from follow-up research certainly indicates that undercontrolled problems show greater stability and poorer long-term prognosis than do overcontrolled problems (e.g., see Esser, Schmidt, & Woerner, 1990; Offord et al., 1992; Robins & Price, 1991). Such findings, however, concern the temporal course of problems independent of therapeutic intervention. What our findings suggest is that when natural time course is held constant through the use of treatment-control comparisons, therapy may not be reliably less effective with undercontrolled problems than with overcontrolled problems.

Although the general type of child problem being treated did not relate to magnitude of treatment outcome, some of our findings suggested that other child characteristics might. Treatment outcomes were better for adolescents than for children. This finding cannot be considered robust, however; the main effect became nonsignificant when we controlled for therapist training. Moreover, Casey and Berman (1985) found no relation between age and therapy outcome, and Weisz et al. (1987) found larger effect sizes for children than for adolescents (i.e., precisely the opposite of the present finding). Some light was shed on this puzzling array of findings by the significant interaction between age and year of publication found in our primary, WLS analysis: The mean effect size for adolescents was significantly greater in studies published between 1986 and 1993 than in pre-1986 studies. This suggests that changes in the age effect from earlier to later meta-analyses may reflect improvements in the efficacy of interventions for adolescents.

In comparison with our age effects, the effects of child gender were more consistent with previous findings. In the present sample of studies, therapy had more beneficial effects in samples with female majorities than in male-majority samples. An important qualification must be added: In the present analysis, the gender effect was highly significant among adolescents but not significant among children, and adolescents showed more

positive treatment effects than children but only in female samples. Thus, overall, psychotherapy showed more beneficial effects for adolescent female samples than for any other Age  $\times$  Gender group. Weisz et al. (1987) did not find the gender main effect found here. In this connection, note that our analyses of year-of-publication effects showed that only in studies published since the 1985 cutoff for the Weisz et al. (1987) sample were girls and adolescents found to be more effectively treated than boys and children. This suggests that secular trends in therapy effects as a function of gender and age may help to explain the difference between the present findings and those of Weisz et al. (1987). Several writers (e.g., Lamb, 1986; Ponton, 1993) have argued that adolescent girls are particularly sophisticated in the use of interpersonal relationships for self-discovery and change and that these skills may facilitate use of the therapeutic relationship to achieve treatment gains. This goodness of fit, along with the fact that therapists are more often female than male, may enhance the impact of treatment for adolescent girls. On the other hand, it is not clear why such an enhancing effect might only be seen in post-1985 studies. It is possible that, since the mid-1980s, interventions have become especially sensitive to the characteristics or treatment needs of adolescent girls, but we have no well-informed opinion as to what specific changes may have been relevant.

Do more fully trained clinicians produce the most beneficial therapy effects? Possibly, but our evidence does not support such a conclusion. Instead, we found that paraprofessionals (typically parents or teachers trained in specific intervention methods) generated larger treatment effects than either student therapists or fully trained professionals; moreover, students and professionals did not differ reliably (this main effect of therapist training was reduced to marginal significance when we controlled for therapy type and for problem type). Such findings are consistent with a growing body of related evidence on intervention effects with children and adults (see Christensen & Jacobson, 1994). We must emphasize, however, that the beneficial effects produced by paraprofessionals and students in these studies followed training and supervision provided by professionals who had, in most cases, designed the procedures. Furthermore, the procedures used may often have been fitted to the training level of therapists; for example, children and procedures assigned to paraprofessionals and students were quite possibly those thought especially appropriate for therapists with little previous training. And, conversely, professionals may be more likely to take on the more difficult, intractable cases.

Thus, the main effect examined here is certainly not a definitive test of the value of professional training. Moreover, we did find a Training  $\times$  Problem Type interaction echoing one found in the Weisz et al. (1987) meta-analysis and pointing to a possible benefit of training: Professional therapists were no more effective than others when treating undercontrolled problems, but professionals produced larger effects than students ( $p < .10$ ) and paraprofessionals ( $p < .001$ ) when treating overcontrolled problems. It is certainly possible that the kinds of behavior management interventions often used with undercontrolled problems tend to be clear cut enough to be taught efficiently to parents and teachers through a focused training program but that the interventions needed for the more subtle and less overt

problems that tend to fall within the overcontrolled category do indeed require substantial professional training.

The interaction between source and domain of outcome measures (reflected in Figure 2) was intriguing and potentially valuable heuristically. For example, the fact that peers, unlike all of the other sources of outcome information, did not perceive positive change in treated children's undercontrolled problems may warrant any of several interpretations; for example, treatment effects with externalizing children may not generalize well to their real-life interactions with peers, or perhaps peers have trouble overcoming reputation effects despite actual behavior change. That trained observers did not report change in children's overcontrolled problems, whereas most other sources did, suggests that direct observation, with its typically brief time samplings and use of observers who do not know the children being observed, may not be a very effective method (in comparison with, for instance, peer, parent, or self-report measures) of assessing outcomes with internalizing youngsters. As a final example, it is intriguing that peers reported very substantial change in treated children's internalizing problems, whereas teachers did not. This may suggest that peers are more sensitive to these rather covert problems than are teachers, whose classroom responsibilities may force them to be more attentive to their students' externalizing problems. Overall, the findings shown in Figure 2 underscore an important fact: One's picture of how a child has responded to treatment is apt to depend on whom one asks (cf. Achenbach, McConaughy, & Howell, 1987).

In addition to its substantive findings, the present study provided a comparison of findings generated by two different analytic methods, the widely used ULS approach (see Smith et al., 1980) and the WLS method (Hedges, 1994; Hedges & Olkin, 1985), which adjusts for heterogeneity of variance across individual studies. In general, it is clear that the WLS approach tends to generate more modest estimates of mean effect size for various groups and (perhaps because of its variance adjustments) more statistically significant differences between groups. Finally, it does not appear that differences between the findings of our 1987 (Weisz et al., 1987) and current meta-analyses can fairly be attributed to our shift to the WLS analytic method because our current WLS findings were no more different from the 1987 findings than were our current ULS findings.

A particularly important focus of our analyses was the question of whether therapy has its impact primarily through non-specific and artifactual effects or through direct impact on targeted problems (see Bowers & Clum, 1988; Frank, 1973; Horvath, 1987, 1988). Our findings showed rather clearly that therapy effects were strongest for outcome measures that specifically matched the problems targeted in treatment. This suggests, in turn, that the therapeutic gains evidenced in these studies were not inadvertent or peripheral effects of some general enhancement in the treated youngsters but were the specific, intended outcomes of interventions targeting specific child problems.

A comparison of the present findings with those of previous child-adolescent meta-analyses reveals some rather consistent patterns (e.g., larger effects for behavioral than nonbehavioral treatments in Casey and Berman, 1985; Weisz et al., 1987, and the present study) and some quite variable patterns (e.g., age effects differ-

ing in direction across meta-analyses). Our evidence suggests that some of the changes in findings (e.g., in gender and age effects) may reflect changes over time in the results of treatment studies; meta-analytic findings change over time because a moving target is being tracked. In addition, however, some of the shifts in findings almost certainly reflect the fact that each individual meta-analysis captures only part of the picture. Thus, it is important that one attend to the cumulative record of successive meta-analyses and regard each individual meta-analytic finding as preliminary until it has been replicated.

As the cumulative record of successive meta-analyses is tracked, it should also be recognized that the findings of most such analyses may be most fairly interpreted as evidence on the state of knowledge about laboratory-based intervention. Relatively few of the studies included in most meta-analyses pertain to psychotherapy as it is practiced in most child and adolescent clinical settings, and the findings may thus reveal little about the effectiveness of most clinic-based interventions (see Weisz, Donenberg, Han, & Kauneckis, in press; Weisz & Weiss, 1993). Clearly, research of the type reviewed here needs to be complemented by research on the impact of intervention with clinic-referred children in service-oriented treatment settings (see Weisz, Donenberg, Han, & Weiss, in press).

Within their proper interpretive context, however, the present findings add substantially to what previous evidence has shown about therapy outcome with young people. Particularly important are our findings that (a) psychotherapy effects were beneficial but weaker than had been thought previously; (b) effects differed as a function of methods of intervention, with behavioral methods showing markedly stronger effects than nonbehavioral approaches; (c) outcomes were related to the interaction of child age and gender, with treatment effects particularly positive for adolescent girls; (d) degree of improvement was a function of the interaction of domain and source of outcome assessment, with improvement in overcontrolled versus undercontrolled behavior depending on the source of information one relies on (e.g., peers vs. parents vs. teachers); and (e) treatment had its strongest effects on the problems targeted in treatment. Continued inquiry will be needed to identify the best analytic methods for estimating effects, to fill out the picture of therapy outcomes across a broad age range, and to sharpen the understanding of factors that can undermine or enhance intervention effects.

## References

References marked with an asterisk indicate studies included in the meta-analysis.

- Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin*, *85*, 1275-1301.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213-232.
- \*Alvord, M. K., & O'Leary, D. K. (1985). Teaching children to share through stories. *Psychology in the Schools*, *22*, 323-330.
- \*Amerikaner, M., & Summerlin, M. L. (1982). Group counseling with learning disabled children: Effects of social skills and relaxation training on self-concept and classroom behavior. *Journal of Learning Disabilities*, *15*, 340-343.
- \*Andrews, D., & Krantz, M. (1982). The effects of reinforced cooperative experience on friendship patterns of preschool children. *Journal of Genetic Psychology*, *140*, 197-205.
- Appelbaum, M. I., & Cramer, E. M. (1974). Some problems in the non-orthogonal analysis of variance. *Psychological Bulletin*, *81*, 335-343.
- \*Arbuthnot, J., & Gordon, D. A. (1986). Behavioral and cognitive effects of a moral reasoning development intervention for high-risk behavior-disordered adolescents. *Journal of Consulting and Clinical Psychology*, *54*, 208-216.
- \*Aury, L. B., & Langenbach, M. (1985). Locus of control and self-responsibility for behavior. *Journal of Educational Research*, *79*, 76-84.
- \*Aydin, G. (1988). The remediation of children's helpless explanatory style and related unpopularity. *Cognitive Therapy and Research*, *12*, 155-165.
- \*Bennett, G. A., Walkden, V. J., Curtis, R. H., Burns, L. E., Rees, J., Gosling, J. A., & McQuire, N. L. (1985). Pad-and-buzzer training, dry-bed training, and stop-start training in the treatment of primary nocturnal enuresis. *Behavioural Psychotherapy*, *13*, 309-319.
- Berman, J. S., Miller, R. C., & Massman, P. J. (1985). Cognitive therapy versus systematic desensitization: Is one treatment superior? *Psychological Bulletin*, *97*, 451-461.
- \*Bierman, K. L., Miller, C. L., & Stabb, S. D. (1987). Improving the social behavior and peer acceptance of rejected boys: Effects of social skill training with instructions and prohibitions. *Journal of Consulting and Clinical Psychology*, *55*, 194-200.
- \*Bloomquist, M. L., August, G. J., & Ostrander, R. (1991). Effects of a school-based cognitive-behavioral intervention for ADHD children. *Journal of Abnormal Child Psychology*, *19*, 591-605.
- \*Blue, S. W., Madsen, C. H., Jr., & Heimberg, R. G. (1981). Increasing coping behavior in children with aggressive behavior: Evaluation of the relative efficacy of the components of a treatment package. *Child Behavior Therapy*, *3*, 51-60.
- \*Bollard, J., Nettelbeck, T., & Roxbee, L. (1982). Dry-bed training for childhood bedwetting: A comparison of group with individually administered parent instruction. *Behaviour Research and Therapy*, *20*, 209-217.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley-Interscience.
- Bowers, T. G., & Clum, G. A. (1988). Relative contribution of specific and nonspecific treatment effects: Meta-analysis of placebo-controlled behavior therapy research. *Psychological Bulletin*, *103*, 315-323.
- Brown, J. (1987). A review of meta-analyses conducted on outcome research. *Clinical Psychology Review*, *7*, 1-23.
- \*Brown, R. T., Borden, K. A., Wynne, M. E., Schleser, R., & Clingerman, S. R. (1986). Methylphenidate and cognitive therapy with ADD children: A methodological reconsideration. *Journal of Abnormal Child Psychology*, *14*, 481-497.
- \*Brown, R. T., Wynne, M. E., & Medenis, R. (1985). Methylphenidate and cognitive therapy: A comparison of treatment approaches with hyperactive boys. *Journal of Abnormal Child Psychology*, *13*, 69-87.
- \*Butler, L., Mieziotis, S., Friedman, R., & Cole, E. (1980). The effect of the two school-based intervention programs on depressive symptoms in preadolescents. *American Journal of Educational Research*, *17*, 111-119.
- \*Campbell, C., & Myrick, R. D. (1990). Motivational group counseling for low-performing students. *Journal for Specialists in Group Work*, *15*, 43-50.
- \*Campbell, D. S., Neill, J., & Dudley, P. (1989). Computer-aided self-instruction training with hearing-impaired impulsive students. *American Annals of the Deaf*, *134*, 227-231.

- Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin*, *98*, 388–400.
- Christensen, A., & Jacobson, N. (1994). Who (or what) can do psychotherapy? The status and challenge of nonprofessional therapies. *Psychological Science*, *5*, 8–14.
- Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.
- \*Costantino, G., Malgady, R. G., & Rogler, L. H. (1986). Cuento therapy: A culturally sensitive modality for Puerto Rican children. *Journal of Consulting and Clinical Psychology*, *54*, 639–645.
- \*Cubberly, W., Weinstein, C., & Cubberly, R. D. (1986). The interactive effects of cognitive learning strategy training and test anxiety on paired-associate learning. *Journal of Educational Research*, *79*, 163–168.
- \*Davidson, W. S., Redner, R., Blakely, C. H., Mitchell, C. M., & Emshoff, J. G. (1987). Diversion of juvenile offenders: An experimental comparison. *Journal of Consulting and Clinical Psychology*, *55*, 68–75.
- \*De Fries, Z., Jenkins, S., & Williams, E. C. (1964). Treatment of disturbed children in foster care. *American Journal of Orthopsychiatry*, *34*, 615–624.
- \*Denkowski, K. M., & Denkowski, G. C. (1984). Is group progressive relaxation training as effective with hyperactive children as individual EMG biofeedback treatment? *Biofeedback and Self-Regulation*, *9*, 353–365.
- \*Downing, C. J. (1977). Teaching children behavior change techniques. *Elementary School Guidance and Counseling*, *12*, 277–283.
- \*Dubey, D. R., O'Leary, S. G., & Kaufman, K. F. (1983). Training parents of hyperactive children in child management: A comparative outcome study. *Journal of Abnormal Child Psychology*, *11*, 229–246.
- \*Dubow, E. F., Huesmann, L. R., & Eron, L. D. (1987). Mitigating aggression and promoting prosocial behavior in aggressive elementary school boys. *Behaviour Research and Therapy*, *25*, 527–531.
- Durlak, J. A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin*, *110*, 204–214.
- \*Edleson, J. L., & Rose, S. D. (1982). Investigations into the efficacy of short-term group social skills training for socially isolated children. *Child Behavior Therapy*, *3*(2/3), 1–16.
- \*Elkin, J. I. W., Weissberg, R. P., & Cowen, E. L. (1988). Evaluation of a planned short-term intervention for schoolchildren with focal adjustment problems. *Journal of Clinical Child Psychology*, *17*, 106–115.
- \*Epstein, L. H., Wing, R. R., Woodall, K., Penner, B. C., Dress, M. J., & Koeske, R. (1985). Effects of a family-based behavioral treatment on obese 5-to-8 year old children. *Behavior Therapy*, *16*, 205–212.
- Esser, G., Schmidt, M. H., & Woerner, W. (1990). Epidemiology and course of psychiatric disorders in school-age children—Results of a longitudinal study. *Journal of Child Psychology and Psychiatry*, *31*, 243–263.
- \*Esters, P., & Levant, R. F. (1983). The effects of two parent counseling programs on rural low-achieving children. *The School Counselor*, *30*, 159–166.
- \*Fehlings, D. L., Roberts, W., Humphries, T., & Dawe, G. (1991). Attention deficit hyperactivity disorder: Does cognitive behavior therapy improve home behavior? *Developmental and Behavioral Pediatrics*, *12*, 223–228.
- \*Fellbaum, G. A., Daly, R. M., Forrest, P., & Holland, C. J. (1985). Community implications of a home-based change program for children. *Journal of Community Psychology*, *13*, 67–74.
- \*Fentress, D. W., Masek, B. J., Mehegan, J. E., & Benson, H. (1986). Biofeedback and relaxation-response training in the treatment of pediatric migraine. *Developmental Medicine and Child Neurology*, *28*, 139–146.
- \*Finney, J. W., Lemanek, K. L., Cataldo, M. F., Katz, H. P., & Fuqua, R. W. (1989). Pediatric psychology in primary health care: Brief targeted therapy for recurrent abdominal pain. *Behavior Therapy*, *20*, 283–291.
- \*Fling, S., & Black, P. (1984/1985). Relaxation covert rehearsal for adaptive functioning in fourth-grade children. *Psychology and Human Development*, *1*, 113–123.
- \*Fox, J. E., & Houston, B. K. (1981). Efficacy of self-instructional training for reducing children's anxiety in an evaluative situation. *Behaviour Research and Therapy*, *19*, 509–515.
- Frank, J. D. (1973). *Persuasion and healing: A comparative study of psychotherapy*. Baltimore: Johns Hopkins University Press.
- \*Friman, P. C., & Leibowitz, M. (1990). An effective and acceptable treatment alternative for chronic thumb- and finger-sucking. *Journal of Pediatric Psychology*, *15*, 57–65.
- \*Gerler, E. R., Jr., Vacc, N. A., & Greenleaf, S. M. (1980). Relaxation training and covert positive reinforcement with elementary school children. *Elementary School Guidance and Counseling*, *15*, 232–235.
- Glass, G. V., & Kliegl, R. M. (1983). An apology for research integration in the study of psychotherapy. *Journal of Consulting and Clinical Psychology*, *51*, 28–41.
- \*Graybill, D., Jamison, M., & Swerdlik, M. E. (1984). Remediation of impulsivity in learning disabled children by special education resource teachers using verbal self-instruction. *Psychology in the Schools*, *21*, 252–254.
- \*Grindler, M. (1988). Effects of cognitive monitoring strategies on the test anxieties of elementary students. *Psychology in the Schools*, *25*, 428–436.
- \*Guerra, N. G., & Slaby, R. G. (1990). Cognitive mediators of aggression in adolescent offenders: 2. Intervention. *Developmental Psychology*, *26*, 269–277.
- \*Gurney, P. W. (1987). Enhancing self-esteem by the use of behavior modification techniques. *Contemporary Educational Psychology*, *12*, 30–41.
- \*Gwynn, C. A., & Brantley, H. T. (1987). Effects of a divorce group intervention for elementary school children. *Psychology in the Schools*, *24*, 161–164.
- \*Hall, J. A., & Rose, S. D. (1987). Evaluation of parent training in groups for parent-adolescent conflict. *Social Work Research and Abstracts*, *23*, 3–8.
- \*Hamilton, S. B., & MacQuiddy, S. L. (1984). Self-administered behavioral parent training: Enhancement of treatment efficacy using a time-out signal seat. *Journal of Clinical Child Psychology*, *13*, 61–69.
- \*Hawkins, J. D., Jenson, J. M., Cataloano, R. F., & Wells, E. A. (1991). Effects of a skills training intervention with juvenile delinquents. *Research on Social Work Practice*, *1*, 107–121.
- \*Hayes, E. J., Cunningham, G. K., & Robinson, J. B. (1977). Counseling focus: Are parents necessary? *Elementary School Guidance and Counseling*, *12*, 8–14.
- Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, *92*, 490–499.
- Hedges, L. V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 285–300). New York: Russell Sage Foundation.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- \*Hiebert, B., Kirby, B., & Jaknavorian, A. (1989). School-based relaxation: Attempting primary prevention. *Canadian Journal of Counseling*, *23*, 273–287.
- \*Hobbs, S. A., Walle, D. L., & Caldwell, H. S. (1984). Maternal evaluation of social reinforcement and time-out: Effects of brief parent training. *Journal of Consulting and Clinical Psychology*, *52*, 135–136.

- Holmbeck, G. N., & Kendall, P. C. (1991). Clinical-childhood-developmental interface: Implications for treatment. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 73-99). New York: Pergamon Press.
- \*Horn, W. F., Ialongo, N. S., Pascoe, J. M., Greenberg, G., Packard, T., Lopez, M., Wagner, A., & Puttler, L. (1991). Additive effects of psychostimulants, parent training, and self-control therapy with ADHD children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 233-240.
- Horvath, P. (1987). Demonstrating therapeutic validity versus the false placebo-therapy distinction. *Psychotherapy*, 24, 47-51.
- Horvath, P. (1988). Placebos and common factors in two decades of psychotherapy research. *Psychological Bulletin*, 104, 214-225.
- \*Huey, W. C., & Rank, R. C. (1984). Effects of counselor and peer-led group assertive training on Black adolescent aggression. *Journal of Counseling Psychology*, 31, 95-98.
- \*Hughes, R. C., & Wilson, P. H. (1988). Behavioral parent training: Contingency management versus communication skills training with or without the participation of the child. *Child and Family Behavior Therapy*, 10(4), 11-23.
- \*Israel, A. C., Stolmaker, L., & Andrian, C. A. G. (1985). The effects of training parents in general child management skills on a behavioral weight loss program for children. *Behavior Therapy*, 16, 169-180.
- \*Jackson, R. L., & Beers, P. A. (1988). Focused videotape feedback psychotherapy: An integrated treatment for emotionally and behaviorally disturbed children. *Counseling Psychology Quarterly*, 1, 11-23.
- \*James, M. R. (1990). Adolescent values clarification: A positive influence on perceived locus of control. *Journal of Alcohol and Drug Education*, 35(2), 75-80.
- \*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.
- \*Jupp, J. J., & Griffiths, M. D. (1990). Self-concept changes in shy, socially isolated adolescents following social skills training emphasizing role plays. *Australian Psychologist*, 25, 165-177.
- Kagan, J. (1965). Impulsive and reflective children: The significance of conceptual tempo. In J. D. Krumboltz (Ed.), *Learning and the educational process*. Chicago: Rand McNally.
- \*Kahn, J. S., Kehle, T. J., Jenson, W. R., & Clark, E. (1990). Comparison of cognitive-behavioral, relaxation, and self-modeling interventions for depression among middle-school students. *School Psychology Review*, 19, 196-211.
- \*Kanigsberg, J. S., & Levant, R. F. (1988). Parental attitudes and children's self-concept and behavior following parents' participation in parent training groups. *Journal of Community Psychology*, 16, 152-160.
- \*Karoly, P., & Rosenthal, M. (1977). Training parents in behavior modification: Effects on perceptions of family interaction and deviant child behavior. *Behavior Therapy*, 8, 406-410.
- 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., Esveltd-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.
- \*Kendall, P. C., & Braswell, L. (1982). Cognitive-behavioral self-control therapy for children: A components analysis. *Journal of Consulting and Clinical Psychology*, 50, 672-689.
- Kendall, P. C., Lerner, R. M., & Craighead, W. E. (1984). Human development and intervention in childhood psychopathology. *Child Development*, 55, 71-82.
- \*Kern, R. M., & Hankins, G. (1977). Adlerian group counseling with contracted homework. *Elementary School Guidance and Counseling*, 12, 284-290.
- \*Kirschenbaum, D. S., Harris, E. S., & Tomarken, A. J. (1984). Effects of parental involvement in behavioral weight loss therapy for preadolescents. *Behavior Therapy*, 15, 485-500.
- \*Klarreich, S. H. (1981). A study comparing two treatment approaches with adolescent probationers. *Corrective and Social Psychiatry and Journal of Behavior Technology, Methods and Therapy*, 27(1), 1-13.
- \*Klein, S. A., & Deffenbacher, J. I. (1977). Relaxation and exercise for hyperactive impulsive children. *Perceptual and Motor Skills*, 45, 1159-1162.
- \*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.
- \*Kotses, H., Harver, A., Segreto, J., Glaus, K. D., Creer, T. L., & Young, G. A. (1991). Long-term effects of biofeedback-induced facial relaxation on measures of asthma severity in children. *Biofeedback and Self-Regulation*, 16, 1-21.
- Lamb, D. (1986). *Psychotherapy with adolescent girls*. New York: Plenum Press.
- \*Larson, K. A., & Gerber, M. M. (1987). Effects of social metacognitive training for enhancing overt behavior in learning disabled and low achieving delinquents. *Exceptional Children*, 54, 201-211.
- \*Larsson, B., Daleflod, B., Hakansson, L., & Melin, L. (1987). Therapist-assisted versus self-help relaxation treatment of chronic headaches in adolescents: A school-based intervention. *Journal of Child Psychology and Psychiatry*, 28, 127-136.
- \*Larsson, B., & Melin, L. (1986). Chronic headaches in adolescents: Treatment in a school setting with relaxation training as compared with information-contact and self-registration. *Pain*, 25, 325-336.
- \*Larsson, B., Melin, L., & Doberl, A. (1990). Recurrent tension headache in adolescents treated with self-help relaxation training and a muscle relaxant drug. *Headache*, 30, 665-671.
- \*Larsson, B., Melin, L., Lamminen, M., & Ullstedt, F. (1987). A school-based treatment of chronic headaches in adolescents. *Journal of Pediatric Psychology*, 12, 553-566.
- \*Lee, D. Y., Hallberg, E. T., & Hassard, H. (1979). Effects of assertion training on aggressive behavior of adolescents. *Journal of Counseling Psychology*, 26, 459-461.
- \*Lee, R., & Haynes, N. M. (1976). Counseling juvenile offenders: An experimental evaluation of Project CREST. *Community Mental Health Journal*, 14, 267-271.
- \*Lewinsohn, P. M., Clarke, G. N., Hops, H., & Andrews, J. (1990). Cognitive-behavioral treatment for depressed adolescents. *Behavior Therapy*, 21, 385-401.
- \*Lewis, R. V. (1983). Scared straight—California style: Evaluation of the San Quentin Squires Program. *Criminal Justice and Behavior*, 10, 209-226.
- \*Liddle, B., & Spence, S. H. (1990). Cognitive-behavior therapy with depressed primary school children: A cautionary note. *Behavioral Psychotherapy*, 18, 85-102.
- \*Lochman, J. E., Burch, P. R., Curry, J. F., & Lampron, L. B. (1984). Treatment and generalization effects of cognitive-behavioral and goal-setting interventions with aggressive boys. *Journal of Consulting and Clinical Psychology*, 52, 915-916.
- \*Lochman, J. E., Lampron, L. B., Gemmer, T. C., Harris, S. R., & Wyckoff, G. M. (1989). Teacher consultation and cognitive-behavioral interventions with aggressive boys. *Psychology in the Schools*, 26, 179-188.
- \*Loffredo, D. A., Omizo, M., & Hammett, V. L. (1984). Group relaxation training and parental involvement with hyperactive boys. *Journal of Learning Disabilities*, 17, 210-213.
- \*Lovaas, O. I. (1987). Behavioral treatment and normal educational

- and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, 55, 3-9.
- \*Lowenstein, L. F. (1982). The treatment of extreme shyness in maladjusted children by impulsive counselling and conditioning approaches. *Acta Psychiatrica Scandinavica*, 66, 173-189.
- Mann, C. (1990). Meta-analysis in the breech. *Science*, 249, 476-480.
- \*Manne, S. L., Redd, W. H., Jacobsen, P. B., Gorfinkle, K., Schorr, O., & Rapkin, B. (1990). Behavioral intervention to reduce child and parent distress during venipuncture. *Journal of Consulting and Clinical Psychology*, 58, 565-572.
- \*McGrath, P. J., Humphreys, P., Goodman, J. T., Keene, D., Firestone, P., Jacob, P., & Cunningham, S. J. (1988). Relaxation prophylaxis for childhood migraine: A randomized placebo-controlled trial. *Developmental Medicine and Child Neurology*, 30, 626-631.
- \*McKenna, J. G., & Gaffney, L. R. (1982). An evaluation of group therapy for adolescents using social skills training. *Current Psychological Research*, 2, 151-160.
- \*McMurray, N. E., Bell, R. J., Fusillo, A. D., Morgan, M., & Wright, F. A. C. (1986). Relationship between locus of control and effects of coping strategies on dental stress in children. *Child and Family Behavior Therapy*, 8(3), 1-15.
- \*McMurray, N. E., Lucas, J. O., Arbes-Duprey, V., & Wright, F. A. C. (1985). The effects of mastery and coping models on dental stress in young children. *Australian Journal of Psychology*, 37, 65-70.
- \*McNeil, C. B., Eyberg, S., Eisenstadt, T. H., Newcomb, K., & Funderburk, B. (1991). Parent-child interaction therapy with behavior problem children: Generalization of treatment effects to the school setting. *Journal of Clinical Child Psychology*, 20, 140-151.
- \*Middleton, H., Zollinger, J., & Keene, R. (1986). Popular peers as change agents for the socially neglected child in the classroom. *Journal of School Psychology*, 24, 343-350.
- \*Miller, D. (1986). Effect of a program of therapeutic discipline on the attitude, attendance, and insight of truant adolescents. *Journal of Experimental Education*, 55, 49-53.
- \*Milne, J., & Spence, S. H. (1987). Training social perception skills with primary school children: A cautionary note. *Behavioural Psychotherapy*, 15, 144-157.
- Mintz, J. (1983). Integrating research evidence: A commentary on meta-analysis. *Journal of Consulting and Clinical Psychology*, 46, 71-75.
- \*Mize, J., & Ladd, G. W. (1990). A cognitive-social learning approach to social skill training with low-status preschool children. *Developmental Psychology*, 26, 388-397.
- \*Moracco, J., & Kazandkian, A. (1977). Effectiveness of behavior counseling and consulting with non-western elementary school children. *Elementary School Guidance and Counseling*, 12, 244-251.
- \*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.
- \*Nall, A. (1973). Alpha training and the hyperkinetic child—Is it effective? *Academic Therapy*, 9, 5-19.
- \*Niles, W. (1986). Effects of a moral development discussion group on delinquent and pre-delinquent boys. *Journal of Counseling Psychology*, 33, 45-51.
- \*Normand, D., & Robert, M. (1990). Modeling of anger/hostility control with preadolescent Type A girls. *Child Study Journal*, 20, 237-261.
- Offord, D. R., Boyle, M. H., Racine, Y. A., Fleming, J. E., Cadman, D. T., Blum, H. M., Byrne, C., Links, P. S., Lipman, E. L., MacMillan, H. L., Grant, N. I. R., Sanford, M. N., Szatmari, P., Thomas, H., & Woodward, C. A. (1992). Outcome, prognosis, and risk in a longitudinal follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31, 916-923.
- \*Oldfield, D. (1986). The effects of the relaxation response on self-concept and acting out behavior. *Elementary School Guidance and Counseling*, 20, 255-260.
- \*Omizo, M. M., & Cubberly, W. E. (1983). The effects of reality therapy classroom meetings on self-concept and locus of control among learning disabled children. *The Exceptional Child*, 30, 201-209.
- \*Omizo, M. M., Hershenberger, J. M., & Omizo, S. A. (1988). Teaching children to cope with anger. *Elementary School Guidance and Counseling*, 22, 241-245.
- \*Omizo, M. M., & Omizo, S. A. (1987). The effects of eliminating self-defeating behavior of learning-disabled children through group counseling. *The School Counselor*, 34, 282-288.
- \*Omizo, M. M., & Omizo, S. A. (1987). The effects of group counseling on classroom behavior and self-concept among elementary school learning disabled children. *The Exceptional Child*, 34, 57-64.
- Parloff, M. B. (1984). Psychotherapy research and its incredible credibility crisis. *Clinical Psychology Review*, 4, 95-109.
- \*Passchier, J., Van Den Bree, M. B. M., Emmen, H. H., Osterhaus, S. O. L., Orelbeke, J. F., & Verhage, F. (1990). Relaxation training in school classes does not reduce headache complaints. *Headache*, 30, 660-664.
- Piaget, J. (1970). Piaget's theory. In P. E. Mussen (Ed.), *Carmichael's manual of child psychology* (3rd ed., Vol. 1, pp. 703-732). New York: Wiley.
- \*Pisterman, S., Firestone, P., McGrath, P., Goodman, J. T., Webster, I., Mallory, R., & Goffin, B. (1992). The role of parent training in treatment of preschoolers with ADHD. *American Journal of Orthopsychiatry*, 62, 397-408.
- \*Pisterman, S., McGrath, P., Firestone, P., Goodman, J. T., Webster, I., & Mallory, R. (1989). Outcome of parent-mediated treatment of preschoolers with attention deficit disorder with hyperactivity. *Journal of Consulting and Clinical Psychology*, 57, 628-635.
- \*Platania-Solazzo, A., Field, T. M., Blank, J., Seligman, F., Kuhn, C., Schanberg, S., & Saab, P. (1992). Relaxation therapy reduces anxiety in child and adolescent psychiatric patients. *Acta Paedopsychiatrica*, 55, 115-120.
- Ponton, L. E. (1993). Issues unique to psychotherapy with adolescent girls. *American Journal of Psychotherapy*, 47, 353-372.
- \*Porter, B. A., & Hoedt, K. C. (1985). Differential effects of an Adlerian counseling approach with preadolescent children. *Individual Psychology: Journal of Adlerian Theory, Research, and Practice*, 41, 372-385.
- \*Potashkin, B. D., & Beckles, N. (1990). Relative efficacy of Ritalin and biofeedback treatments in the management of hyperactivity. *Biofeedback and Self-Regulation*, 15, 305-315.
- \*Rae, W. A., Worchel, F. F., Upchurch, J., Sanner, J. H., & Daniel, C. A. (1989). The psychosocial impact of play on hospitalized children. *Journal of Pediatric Psychology*, 14, 617-627.
- \*Raue, J., & Spence, S. H. (1985). Group versus individual applications of reciprocity training for parent-youth conflict. *Behaviour Research and Therapy*, 23, 177-186.
- \*Redfering, D. L., & Bowman, M. J. (1981). Effects of a meditative-relaxation exercise on non-attending behaviors of behaviorally disturbed children. *Journal of Clinical Child Psychology*, 10, 126-127.
- Reed, J. G., & Baxter, P. M. (1994). Scientific communication and literature retrieval. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 57-70). New York: Russell Sage Foundation.
- \*Reynolds, W. M., & Coats, K. I. (1986). A comparison of cognitive-behavioral therapy and relaxation training for the treatment of depression in adolescents. *Journal of Consulting and Clinical Psychology*, 54, 653-660.
- Rice, F. P. (1984). *The adolescent: Development, relationships, and culture*. Boston: Allyn & Bacon.

- \*Richter, I. L., McGrath, P. J., Humphreys, P. J., Firestone, P., & Keene, D. (1986). Cognitive and relaxation treatment of paediatric migraine. *Pain*, 25, 195-203.
- \*Rivera, E., & Omizo, M. M. (1980). The effects of relaxation and biofeedback on attention to task and impulsivity among male hyperactive children. *The Exceptional Child*, 27, 41-51.
- Robins, L. N., & Price, R. K. (1991). Adult disorders predicted by childhood conduct problems: Results from the NIMH Epidemiologic Catchment Area Project. *Psychiatry*, 54, 116-132.
- 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.
- \*Rosenfarb, I., & Hayes, S. C. (1984). Social standard setting: The Achilles heel of information accounts of therapeutic change. *Behavior Therapy*, 15, 515-528.
- \*Saigh, P. A., & Antoun, F. T. (1984). Endemic images and the desensitization process. *Journal of School Psychology*, 22, 177-183.
- \*Sanders, M. R., Rebetz, M., Morrison, M., Bor, W., Gordon, A., Dadds, M., & Shepherd, R. (1989). Cognitive-behavioral treatment of recurrent nonspecific abdominal pain in children: An analysis of generalization, maintenance, and side effects. *Journal of Consulting and Clinical Psychology*, 57, 294-300.
- \*Sarason, I. G., & Sarason, B. R. (1981). Teaching cognitive and social skills to high school students. *Journal of Consulting and Clinical Psychology*, 49, 908-918.
- \*Schneider, B. H., & Byrne, B. M. (1987). Individualizing social skills training for behavior-disordered children. *Journal of Consulting and Clinical Psychology*, 55, 444-445.
- \*Scott, M. J., & Stradling, S. G. (1987). Evaluation of a group programme for parents of problem children. *Behavioural Psychotherapy*, 15, 224-239.
- \*Senediak, C., & Spence, S. H. (1985). Rapid versus gradual scheduling of therapeutic contact in a family based behavioural weight control programme for children. *Behavioural Psychotherapy*, 13, 265-287.
- \*Seymour, F. W., Brock, P., Doring, M., & Poole, G. (1989). Reducing sleep disruptions in young children: Evaluation of therapist-guided and written information approaches: A brief report. *Journal of Child Psychology and Psychiatry*, 30, 913-918.
- \*Shechtman, Z. (1991). Small group therapy and preadolescent same-sex friendship. *International Journal of Group Psychotherapy*, 41, 227-243.
- \*Sherer, M. (1985). Effects of group intervention on moral development of distressed youths in Israel. *Journal of Youth and Adolescence*, 14, 513-526.
- Shirk, S. (Ed.). (1988). *Cognitive development and child psychotherapy*. New York: Plenum Press.
- \*Skuy, M., & Solomon, W. (1980). Effectiveness of students and parents in a psycho-educational intervention programme for children with learning problems. *British Journal of Guidance and Counselling*, 8, 76-86.
- \*Sluckin, A., Foreman, N., & Herbert, M. (1991). Behavioural treatment programs and selectivity of speaking at follow-up in a sample of 25 selective mutes. *Australian Psychologist*, 26, 132-137.
- Smith, M. L., Glass, G. V., & Miller, T. L. (1980). *The benefits of psychotherapy*. Baltimore: Johns Hopkins University Press.
- \*Smyrniotis, K. X., & Kirkby, R. J. (1993). Long-term comparison of brief versus unlimited psychodynamic treatments with children and their parents. *Journal of Consulting and Clinical Psychology*, 61, 1020-1027.
- \*Spaccarelli, S., & Penman, D. (1992). Problem-solving skills training as a supplement to behavioral parent training. *Cognitive Therapy and Research*, 16, 1-18.
- \*Stark, K. D., Reynolds, W. M., & Kaslow, N. J. (1987). A comparison of the relative efficacy of self-control therapy and a behavioral problem-solving therapy for depression in children. *Journal of Abnormal Child Psychology*, 15, 91-113.
- \*Steele, K., & Barling, J. (1982). Self-instruction and learning disabilities: Maintenance, generalization, and subject characteristics. *Journal of General Psychology*, 106, 141-154.
- \*Stevens, R., & Pihl, R. O. (1982). The remediation of the student at risk for failure. *Journal of Clinical Psychology*, 38, 298-301.
- \*Stevens, R., & Pihl, R. O. (1983). Learning to cope with school: A study of the effects of a coping skill training program with test-vulnerable 7th-grade students. *Cognitive Therapy and Research*, 7, 155-158.
- \*Stewart, M. J., Vockell, E., & Ray, R. E. (1986). Decreasing court appearances of juvenile status offenders. *Social Casework: The Journal of Contemporary Social Work*, 67, 74-79.
- \*Strayhorn, J. M., & Weidman, C. S. (1989). Reduction of attention deficit and internalizing symptoms in preschoolers through parent-child interaction training. *Journal of the American Academy of Child and Adolescent Psychiatry*, 28, 888-896.
- \*Strayhorn, J. M., & Weidman, C. S. (1991). Follow-up one year after parent-child interaction training: Effects on behavior of preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31, 138-143.
- \*Sutton, C. (1992). Training parents to manage difficult children: A comparison of methods. *Behavioural Psychotherapy*, 20, 115-139.
- \*Szapocznik, J., Rio, A., Murray, E., Cohen, R., Scopetta, M., Rivas-Vazquez, A., Hervis, O., Posada, V., & Kurtines, W. (1989). Structural family versus psychodynamic child therapy for problematic Hispanic boys. *Journal of Consulting and Clinical Psychology*, 57, 571-578.
- \*Tanner, V. L., & Holliman, W. B. (1988). Effectiveness of assertiveness training in modifying aggressive behaviors of young children. *Psychological Reports*, 62, 39-46.
- \*Thacker, V. J. (1989). The effects of interpersonal problem-solving training with maladjusted boys. *School Psychology International*, 10, 83-93.
- \*Trapani, C., & Gettinger, M. (1989). Effects of social skills training and cross-age tutoring on academic achievement and social behaviors of boys with learning disabilities. *Journal of Research and Development in Education*, 22(4), 1-9.
- \*Van Der Ploeg, H. M., & Van Der Ploeg-Stapert, J. D. (1986). A multifaceted behavioral group treatment for test anxiety. *Psychological Reports*, 58, 535-542.
- \*Vaughn, S., & Lancelotta, G. X. (1990). Teaching interpersonal social skills to poorly accepted students: Peer-pairing versus non-peer-pairing. *Journal of School Psychology*, 28, 181-188.
- \*Verduyn, C. M., Lord, M., & Forrest, G. C. (1990). Social skills training in schools: An evaluation study. *Journal of Adolescence*, 13, 3-16.
- \*Warren, R., Smith, G., & Velten, E. (1984). Rational-emotive therapy and the reduction of interpersonal anxiety in junior high school students. *Adolescence*, 29, 893-902.
- \*Webster-Stratton, C. (1992). Individually administered videotape parent training: "Who benefits?" *Cognitive Therapy and Research*, 16, 31-35.
- \*Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting and Clinical Psychology*, 56, 558-566.
- \*Wehr, S. H., & Kaufman, M. E. (1987). The effects of assertive training on performance in highly anxious adolescents. *Adolescence*, 85, 195-205.
- Weiss, B., & Weisz, J. R. (1990). The impact of methodological factors on child psychotherapy outcome research: A meta-analysis for researchers. *Journal of Abnormal Child Psychology*, 18, 639-670.

- Weiss, B., & Weisz, J. R. (in press). Relative effectiveness of behavioral versus nonbehavioral child psychotherapy. *Journal of Consulting and Clinical Psychology*.
- Weisz, J. R., Donenberg, G. R., Han, S. S., & Kauneckis, D. (in press). Child and adolescent psychotherapy outcomes in experiments versus clinics: Why the disparity? *Journal of Abnormal Child Psychology*.
- Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (in press). Bridging the gap between lab and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*.
- Weisz, J. R., & Weiss, B. (1993). *Effects of psychotherapy with children and adolescents*. Newbury Park, CA: 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.
- White, H. D. (1994). Scientific communication and literature retrieval. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 41-56). New York: Russell Sage Foundation.
- \*Wilson, N. H., & Rotter, J. C. (1986). Anxiety management training and study skills counseling for students on self-esteem and test anxiety and performance. *The School Counselor*, 34, 18-31.
- \*Wisniewski, J. J., Genshaft, J. L., Mulick, J. A., Coury, D. L., & Hammer, D. (1988). Relaxation therapy and compliance in the treatment of adolescent headache. *Headache*, 28, 612-617.
- \*Zieffle, T. H., & Romney, D. M. (1985). Comparison of self-instruction and relaxation training in reducing impulsive and inattentive behavior of learning disabled children on cognitive tasks. *Psychological Reports*, 57, 271-274.
- \*Ziegler, E. W., Scott, B. N., & Taylor, R. L. (1991). Effects of guidance intervention on perceived school behavior and personal self-concept: A preliminary study. *Psychological Reports*, 68, 50.

Received April 25, 1994

Revision received October 27, 1994

Accepted October 27, 1994 ■

### New Journal: *Psychological Methods*

The Publications and Communications Board of the American Psychological Association announces the publication of a new quarterly journal, *Psychological Methods*, beginning in 1996.

Although originating from the *Psychological Bulletin* section on "Quantitative Methods," the new journal has an expanded coverage policy and will be devoted to the development and dissemination of methods for collecting, analyzing, understanding, and interpreting psychological data. Its purpose is the dissemination of innovations in research design, measurement, methodology, and quantitative and qualitative analysis to the psychological community; its further purpose is to promote effective communication about related substantive and methodological issues. The audience is expected to be diverse and to include those who develop new procedures and those who employ those procedures in research, as well as those who are responsible for undergraduate and graduate training in design, measurement, and statistics. The journal solicits original theoretical, quantitative, empirical, and methodological articles; reviews of important methodological issues as well as of philosophical issues with direct methodological relevance; tutorials; articles illustrating innovative applications of new procedures to psychological problems; articles on the teaching of quantitative methods; and reviews of statistical software. Submissions will be judged on their relevance to understanding psychological data, methodological correctness, and accessibility to a wide audience. Where appropriate, submissions should illustrate through concrete example how the procedures described or developed can enhance the quality of psychological research. The journal welcomes submissions that show the relevance to psychology of procedures developed in other fields. Empirical and theoretical articles on specific tests or test construction should have a broad thrust; otherwise, they may be more appropriate for *Psychological Assessment*.

Mark I. Appelbaum, PhD, has been appointed to a 6-year term as editor of *Psychological Methods*. Beginning immediately, manuscripts should be sent to

Mark I. Appelbaum, PhD  
 Department of Psychology and  
 Human Development  
 Box 159 Peabody  
 Vanderbilt University  
 Nashville, TN 37203