Journal of Consulting and Clinical Psychology
1989, Vol. 57, No. 6, 741-746


0022-006X/89/$00.75

# Assessing the Effects of Clinic-Based Psychotherapy With Children and Adolescents

John R. Weisz
University of North Carolina

Bahr Weiss
Emma Pendleton Bradley Hospital
Brown University School of Medicine

Recent meta-analyses suggest that psychotherapy is quite effective with children and adolescents. However, most research in those analyses involved controlled laboratory interventions that may not represent typical therapy in clinics. We studied more representative treatment as it routinely occurs, in 9 clinics. We compared 93 youngsters who completed a course of therapy with 60 who dropped out after intake. At intake, the groups did not differ on demographic, family, or clinical measures, including Child Behavior Checklist (CBCL) scores. Six months later (when therapy had ended for 98% of the treated children) and again 1 year later, the 2 groups were compared on CBCL scores, parent ratings of each child's major referral problem, and (for a subsample) teacher reports. No comparison showed significant main effects of therapy. The findings (a) raise questions about the generalizability of findings from research-oriented therapy and (b) suggest that the control and precision of research therapy may be needed in clinical practice.

How effective is psychotherapy with children and adolescents? Two recent meta-analyses offer quite positive assessments. Casey and Berman (1985) analyzed outcome studies involving children 12 years of age or younger, and Weisz, Weiss, Alicke, and Klotz (1987) analyzed studies with children and adolescents aged 4–18 years. In both analyses, the average treated youngster functioned better after treatment than three fourths of the untreated controls. These positive findings are subject to a potentially important limitation, however: Most of the studies reviewed may have involved conditions and interventions unrepresentative of usual clinical practice. For example, in the large majority of studies in Weisz, Weiss, Alicke, & Klotz (1987), (a) youngsters were recruited for treatment rather than being clinic-referred; (b) samples were selected for homogeneity, with all youngsters displaying a similar problem (e.g., a specific phobia); (c) therapy was focused primarily or exclusively on the focal problem or problems; (d) therapists were specially selected and were trained immediately prior to therapy in the specific techniques they would use; and (e) therapy involved nearly exclusive reliance on those techniques.

In most clinical practice, by contrast, (a) clients are seriously enough disturbed to have been clinic-referred; (b) clients are heterogeneous, with most youngsters referred for multiple problems; (c) therapy is directed not at one focal problem but at a broad spectrum of problems for each youngster; (d) therapists are unlikely to have had recent training in most of the techniques they use; and (e) therapy is not confined to a few techniques in which therapists are well versed. Conditions in most clinics may thus be rather different from those arranged for outcome research. A key question thus arises: Are the benefits of therapy as demonstrated in controlled-outcome studies also found in more typical clinical practice?

One way to answer this question is to study therapy outcomes for children who are spontaneously referred and treated according to usual clinic procedures. However, the gain in external validity offered by such research must be paid for in methodological compromises. Ethical and legal constraints prevent clinics from making truly random assignments of referred youth to no-treatment control groups. Thus, rigorous treatment–control comparisons are seldom feasible in such settings. To cope with this problem, some researchers (e.g., Levitt, 1971; McAdoo & Roeske, 1973; Weisz, Weiss, & Langmeyer, 1987) have proposed that under certain conditions, children who begin treatment but do not continue may be used as a control group to be compared with those who complete a course of therapy.

But might not children who continue in therapy differ from dropouts clinically or demographically? Actually, published research has revealed negligible group differences (Gould, Shaffer, & Kaplan, 1985; Levitt, 1957; Weisz, Weiss, & Langmeyer, 1987); this supports the argument that children who drop out may be an acceptable (though certainly not ideal), naturally occurring control group for outcome research. Building on this notion, we compared dropouts with continuers in an effort to shed early light on the question of real-life therapy effects.

This study was conducted while John R. Weisz was supported by North Carolina Department of Human Resources Grant 41626 and National Institute of Mental Health Grant 1 R03 MH38450.

We are grateful to the children and parents who took part in the study and to staff members of all the participating clinics. We are also grateful to Thomas M. Achenbach and W. Edward Craighead for their thoughtful comments on an earlier draft of this article.

Further details regarding the data and statistical analyses can be obtained from John R. Weisz for $3 to cover costs.

Correspondence concerning this article should be addressed to John R. Weisz, Department of Psychology, Davie Hall CB# 3270, University of North Carolina, Chapel Hill, North Carolina 27599-3270.

## Method

### *Six-Month Follow-Up Groups*

The 6-month subject pool consisted of 262 outpatient clinic-referred 6- to 17-year-olds for whom parents had completed a Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983) at intake and again 6 months later. Following McAdoo and Roeske (1973), we defined *dropouts* as cases in which clinic treatment had been recommended and offered but in which neither the child nor other family members had appeared for any sessions after intake; we excluded 14 children who had received other mental health services during the period of this study. Following McAdoo and Roeske (1973), we defined *continuers* as children who, with or without other family members, continued for at least five sessions and terminated with concurrence of the therapist. Continuers averaged 12.43 sessions ($SD = 5.53$). The 95 children who had two to four sessions or terminated against their therapist's recommendation were excluded. This left 60 dropouts and 93 continuers, seen in nine public outpatient clinics in nine municipalities and assigned to 37 different therapists (who averaged 2.43 subjects each; range = 1–8); this diversity helped ensure that findings would not reflect idiosyncratic characteristics of any single therapist, clinic, or location. These 153 children were a subset of the 304 originally used in a study of the characteristics of children who drop out of therapy (Weisz, Weiss, & Langmeyer, 1987); that study also used the child demographic variables used here, plus the CBCL at the time of intake only. The dropout and continuer groups did not differ significantly on therapist sex, race (Fisher's Exact Test [FET]), or years of experience ($t$ tests).

### *One-Year Follow-Up Groups*

The subject pool for the 1-year follow-up included 246 clinic-referred 6- to 17-year-olds from the same nine clinics. We excluded 11 from the dropout group because they received other services; 94 others were dropped because they had from two to four sessions or terminated against therapist recommendation. This left 69 dropouts and 72 continuers; 34 of the dropouts and 52 continuers were also in the 6-month sample. Continuers averaged 13.10 sessions ($SD = 5.76$) and were seen by 28 different therapists (who averaged 2.46 subjects; range = 1–7). Continuer versus dropout differences were again nonsignificant with respect to therapist sex, race, age, and experience.

### *Potential Predictors of Treatment Effects*

We examined the same potential predictors of therapy effects studied in the Weisz, Weiss, Alicke, and Klotz (1987) meta-analysis: child sex, child age (6–12 years vs. 13–17 years), problem type (CBCL internalizing vs. externalizing; see subsequent details), therapist training (i.e., students [trainees, interns, and child fellows] vs. professionals [clinic staff who had doctoral or master's degrees]), and therapy type. The youngsters primarily received individual sessions. To gauge therapy type, we asked each therapist who worked with a continuer to estimate the percentage of therapy with that child involving psychodynamic, cognitive, behavioral, and other methods. We received estimates for 92% of the 6-month sample and 93% of the 1-year sample. "Cognitive" approaches were generally cognitive–behavioral, so we grouped cognitive and behavioral into a single behavioral category, as in the Weisz, Weiss, Alicke, and Klotz (1987) analysis (cf. Weisz, 1986).

### *Procedure and Measures*

*Outcome Measure 1: CBCL problem reports from parents.* At intake (Time 1), 6 months later (Time 2) when all but 2 of the treated children had finished therapy, and 12 months later (Time 3), the parent completed the CBCL for the referred child. The CBCL surveys a broad spectrum of clinically significant problem behaviors and has proven sensitive to differences between treated and untreated groups in controlled-outcome studies (e.g., Webster-Stratton, 1984). The CBCL yields $T$ scores for internalizing (e.g., sadness and worrying), externalizing (e.g., arguing and fighting), total problems, and social competence.

In soliciting CBCL reports, we sent one or two mailouts, then used phone reminders, until a parent expressed unwillingness to participate. Research ethics then required that we stop; moreover, unwilling parents might have provided data of doubtful validity. Of the parents for whom addresses were known, 55% returned a fully completed CBCL at 6 months and 51% at 1 year. At 6 months, parents of dropouts were less likely to return the CBCL than parents of continuers, $\chi^2(1, N = 304) = 15.84$, $p < 0.001$ (304 is the total of dropouts and continuers for whom we had a CBCL at intake). At 1 year, the two groups did not differ in return rates, ($\chi^2 < 1$).

*Outcome Measure 2: Focal problems identified by parents.* At the end of the CBCL, we asked parents to write up to three "major problems for which your child needs help" and to rate the severity of each problem with a 0, 1, or 2 (same scale as the CBCL). In the 6-month sample, 99 parents did so at Times 1 and 2, whereas in the 1-year sample, 79 parents did so at Times 1 and 3; among parents who returned the CBCL, dropout and continuer percentages were not reliably different in the 6- and 12-month samples (FETs, *ns*).

*Outcome Measure 3: Teacher reports on a subsample.* Parent reports were our primary index of child functioning, but we collected teacher reports on a small subsample. Each time parents filled in the CBCL, we asked them to identify a teacher who could report on their child's current behavior at school, using the CBCL Teacher Report Form (TRF; Achenbach & Edelbrock, 1986). Many parents declined; some did not want their child singled out, and for others, summer vacation meant that there was no current teacher. We contacted teachers who were identified, asking them to take part in a "Carolina Youth Survey" by filling in the TRF for the child (who was not identified as a clinic client). We did not call unusual attention to the child by phone or repeated mail contacts. These constraints led to low sample sizes: 36 in the 6-month sample, 33 in the 1-year sample. At 6 months the participation rate was higher for continuers than dropouts (FET $p = .002$). At 1 year the direction of the group difference was reversed but nonsignificant (FET, *ns*).

### *Assessing Representativeness of Dropout and Continuer Samples*

To check representativeness of the 6- and 12-month samples, we compared participating and nonparticipating dropouts and participating and nonparticipating continuers on 11 demographic, family, and child clinical variables assessed at Time 1. To minimize the probability of a Type II error, we made no adjustment for the number of tests at this point or later in the comparison of dropouts and continuers. In the 6-month sample, none of the 11 tests comparing participating and nonparticipating dropouts was significant (all $ps > .05$): child age, sex, birth order, number of children at home, miles to the clinic, changes in family structure (e.g., parent separation), Children's Depression Inventory (CDI) score, number of previous outpatient sessions, CBCL $T$ scores for internalizing and externalizing problems, and social competence. Of the 11 tests comparing participating and nonparticipating continuers, one was significant: Changes in family structure averaged 1.78 for nonparticipators and 1.92 for participators, $t(136) = -2.50$, $p < .05$. All other tests were nonsignificant (all $ps > .05$): child age, sex, birth order, number of children at home, miles to the clinic, CDI score, previous outpatient sessions, and internalizing, externalizing, and social competence scores.

Table 1
*Comparison of Dropout and Continuer Groups at Time 1*

| | 6-month sample | | | | 1-year sample | | | |
|---|---|---|---|---|---|---|---|---|
| | Dropout | | Continuer | | Dropout | | Continuer | |
| Variable | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Age at intake (in years) | 10.88 | 3.15 | 11.00 | 3.01 | 11.39 | 3.25 | 11.03 | 3.17 |
| Hollingshead SES (9-point scale) | 4.22 | 1.95 | 4.48 | 1.98 | 4.43 | 1.99 | 4.46 | 1.95 |
| CBCL total problems *T* | 70.40 | 13.79 | 71.44 | 9.41 | 70.33 | 11.68 | 71.04 | 9.04 |
| CBCL social competence *T* | 36.00 | 9.12 | 35.67 | 8.49 | 35.30 | 8.92 | 35.14 | 8.28 |
| No. children living at home | 2.39 | 1.75 | 2.06 | 1.15 | 2.05 | 1.26 | 2.16 | 1.18 |
| Miles from home to clinic | 3.00 | 5.81 | 4.10 | 6.86 | 3.63 | 5.83 | 3.66 | 7.07 |
| Changes in family | 1.91 | 0.29 | 1.92 | 0.27 | 1.88 | 0.33 | 1.92 | 0.28 |
| CDI score (depression) | 12.26 | 7.88 | 11.33 | 6.92 | 12.65 | 8.28 | 11.92 | 7.17 |
| Number previous sessions | 0.33 | 0.82 | 0.49 | 1.85 | 0.31 | 0.78 | 0.56 | 2.08 |
| % male | 63.33 | | 64.52 | | 59.42 | | 76.39 | |

*Note.* SES = socioeconomic status; CBCL = Child Behavior Checklist; *T* = *T* score; CDI = Children's Depression Inventory.

In the 1-year sample, none of the 11 tests for dropouts was significant (all *ps* > .15): child age, sex, birth order, number of children at home, miles to the clinic, changes in family structure, CDI score, previous outpatient sessions, and internalizing, externalizing, and social competence scores. Of the 11 tests for continuers, one was significant: Participators were 76% male, and nonparticipators were 48% male (FET $p = .0009$). All other tests were nonsignificant (all *ps* > .10): child age, birth order, number of children at home, miles to the clinic, changes in family structure, CDI score, previous outpatient sessions, and internalizing, externalizing, and social competence scores. In 44 tests, the two significant differences fall well within chance expectancy (Feild & Armenakis, 1974). So, the dropout and continuer samples appear representative of the overall pool of dropouts and continuers.

### *Assessing Similarity of Dropouts and Continuers*

To assess appropriateness of posttreatment comparisons between the dropouts and continuers, we tested whether the groups were similar on measures that might be related to improvement. We compared the groups in the 6-month and 1-year samples on the same 11 variables mentioned earlier. In the 6-month sample, none of the 11 tests were significant (all *ps* > .20; see Table 1 for means): child age, sex, birth order, number of children at home, miles to the clinic, changes in family structure, CDI score, previous outpatient sessions, and internalizing, externalizing, and social competence scores.

In the 1-year sample, the two groups differed significantly in regard to sex, with continuers being 76% male and dropouts 59% male (FET $p = .04$). All other tests were nonsignificant (all *ps* > .45): child age, birth order, number of children at home, miles to the clinic, changes in family structure, CDI score, previous outpatient sessions, and internalizing, externalizing, and social competence scores. In 22 tests, the one significant difference is well within chance expectancy and therefore not reliable. Thus, the two groups were not reliably different in the 6-month or the 1-year sample across the 11 comparison variables. It is possible, however, that the groups differed on child or family characteristics not assessed here and that such differences influenced the findings reported hereinafter.

## Results

Our primary analyses focused on changes in CBCL scores from Time 1 to 2 and from Time 1 to 3. Unless otherwise noted, (a) analyses were analyses of covariance (ANCOVAs) with Time 1 score as covariate and Time 2 or 3 score as dependent variable, and (b) the relation between a variable (e.g., age) and improvement was assessed by testing the interaction between that variable and the Group factor (i.e., dropout vs. continuer). To limit Type I error, we used the Bonferroni alpha adjustment procedure (Kirk, 1982), adjusting according to the number of tests in a family (excluding eliminating tests, see details that follow). We did not apply the alpha adjustment to (a) tests of overall therapy effects (because it was most important to avoid Type II error) or to (b) post hoc tests used only to explain higher order effects.

### *Overall Effect of Therapy On Outcome Measure 1: CBCL Scores*

We first estimated overall therapy effects, comparing changes in CBCL total problem scores in dropouts and continuers. A one-way ANCOVA, with Time 1 CBCL score as covariate and group as the independent variable, showed no reliable difference between groups in change from Time 1 to Time 2, $F(1, 150) = 3.32$, $p < .10$, or from Time 1 to Time 3, $F(1, 138) = 2.93$, $p < .10$. As Table 2 shows, both groups improved significantly from Time 1 to Time 2, dropout $t(59) = 3.68$, $p < .001$, continuer $t(92) = 3.09$, $p < .01$, and from Time 1 to Time 3, dropout $t(68) = 3.00$, $p < .01$, continuer $t(71) = 5.89$, $p < .001$. We reran these as well as the other main effect tests of therapy (i.e., those involving focal problems and teacher reports), dropping from the 6-month sample the 2 subjects who had not completed treatment at the time of the second assessment; results were unchanged.

As a rough comparison between findings of this study and the Weisz, Weiss, Alicke, and Klotz (1987) meta-analysis, we calculated the effect size (ES) for the present data. Standard meta-analysis ESs are based on posttreatment means for treatment and control groups, unadjusted for Time 1 scores; group differences at Time 1 are expected to average out across multi-

Table 2
*Posttreatment Scores Adjusted for Pretreatment Scores on the Three Outcome Measures*

| Measure | At 6 months | At 1 year |
|---|---|---|
| CBCL total problems | | |
| Dropouts | 65.56 | 66.95 |
| Continuers | 68.32 | 64.27 |
| Focal problems[a] | | |
| Dropouts | 1.63 | 1.41 |
| Continuers | 1.46 | 1.29 |
| TRF total problems | | |
| Dropouts | 65.00 | 66.61 |
| Continuers | 62.04 | 63.73 |

*Note.* CBCL = Child Behavior Checklist; TRF = Teacher Report Form.
[a] Scale ranges from *not true of the child* (0) to *very true or often true of the child* (2).

ple studies. In a single study, however, this averaging does not occur. So we also computed effect size approximations (ESAs) using posttreatment scores adjusted for Time 1 scores. For both approaches, we subtracted the dropout mean from the continuer mean and divided by the standard deviation of the dropout group. The resulting ESs and ESAs are standard scores; 0 indicates no treatment effect, a positive value a beneficial effect, and a negative value an adverse effect.

The ES for CBCL scores at 6 months was −0.24 ($p < .10$), indicating that the group that completed treatment fared somewhat worse than the group that dropped out. At 1 year, the ES was 0.19 ($p > .25$). The ESA for CBCL scores at 6 months was −0.26 ($p < .10$), again suggesting that continuers fared nonsignificantly less well than did dropouts. At 1 year, the ESA was 0.30 ($p < .10$).

These ES and ESA values contrast with the highly significant ESs of 0.71 and 0.79, respectively, in the Casey and Berman (1985) and Weisz, Weiss, Alicke, and Klotz (1987) meta-analyses. In the latter study, the 95% confidence interval for the posttreatment ES of 0.79 was 0.64–0.94; the 95% confidence interval for the mean ES of 0.93 in follow-ups was 0.55–1.38. Thus, the 6-month and 1-year ESs and ESAs here were well outside the confidence intervals of the meta-analysis effect sizes.

## *Effects of Child, Therapy, and Therapist Factors on CBCL Scores*

We next patterned a series of ANCOVAs after the Weisz, Weiss, Alicke, and Klotz (1987) meta-analysis, testing whether improvement was related to child age or sex, problem type, therapy type, or therapist training (adjusted alpha for each family was 0.01). As in the meta-analysis, cell sizes ruled out inclusion of more than two factors (in addition to Group, and the Time 1 score as covariate) in any analysis. Yet we did want to control for confounding among the factors. So, following Weisz, Weiss, Alicke, and Klotz (1987), we tested (a) main effects of each factor (technically, interactions between the factors of interest and the Group factor), ignoring the other four factors; (b) main effects of each factor, eliminating (i.e., controlling for) the other four factors individually (see Appelbaum & Cramer, 1974); and (c) all possible two-way interactions. In describing the results of interaction tests subsequently, we do not report redundant analyses (e.g., we report the Age × Sex test only in the Child age section). Because problem type was a repeated measure analyzed multivariately, we could not eliminate it in analyses of the four between-subjects factors, nor could we eliminate those four in the problem type analysis. Because therapy type was grouped into three levels—continuer-behavioral, continuer-nonbehavioral, and dropout—it could not be eliminated in tests of other factors. Finally, there were too few student therapists to permit tests of therapist training interactions.

*Child age.* We tested the age effect, comparing treatment–no treatment differences in children (aged 6–12 years) and adolescents (13–17 years). From Time 1 to 2, the age effect was not significant with other factors ignored ($p < .10$), or when eliminating sex, ($p < .10$) or therapist training, $F(1, 80) = 4.36$, $p < .05$. From Time 1 to 3, the age effect was not significant with other factors ignored, $F(1, 136) = 0.00$, $p > .50$), or when eliminating sex, $F(1, 132) = 3.00$, $p > .15$, or therapist training ($p > .50$). We next examined interactions involving age. The Age × Sex interaction was not significant from Time 1 to 2 or from Time 1 to 3. Similarly, neither the Age × Therapy Type nor the Age × Problem Type interaction was significant for the 6- and 12-month samples, respectively.

*Child sex.* From Time 1 to 2 and from 1 to 3, the sex effect was not significant (respectively) with other factors ignored or with age or therapist training eliminated. Interactions of sex with therapy type and problem type were also nonsignificant from Time 1 to 2 and from 1 to 3, respectively.

*Problem type.* The main effect of problem type was not significant from Time 1 to 2 or from 1 to 3. Nor was the interaction of problem type with therapy type significant from Time 1 to 2 or from 1 to 3.

*Therapy type.* We tested effects of therapy type, comparing children who received primarily behavioral treatment, primarily nonbehavioral treatment, or no treatment (i.e., dropouts). From Time 1 to 2 and from 1 to 3, respectively, therapy type was not significant with other factors ignored or with age, sex, or therapist training eliminated.

*Therapist training.* The main effect of therapist training was not significant from Time 1 to 2 or from 1 to 3, respectively, with other factors ignored or with age or sex eliminated.

## *Overall Effect of Therapy on Outcome Measure 2: Focal Problem Ratings*

We next focused on the specific problems parents noted as their child's major problems. We used ANCOVAs to analyze mean ratings across all problems a parent noted. Although in both samples, continuers showed somewhat more change from pre- to posttreatment than dropouts showed (see Table 2), therapy effects were not significant at 6 months or 1 year. For the 6-month sample, the ES was 0.24, and the ESA was 0.32. For the 1-year sample, the ES was 0.24, and the ESA was 0.20.

### *Effects of Child, Therapy, and Therapist Factors on Focal Problems*

We next conducted ignoring, eliminating, and interaction tests (adjusted alpha = 0.01) for child age and sex, therapy type, and therapist training; the tests paralleled those described previously for CBCL scores. Across the 6- and 12-month samples, one test was significant: For the 1-year sample, the effect for sex was significant, with other factors ignored, $F(1, 74) = 7.59$, $p < .01$, as well as with age, $F(1, 70) = 9.20$, $p < .005$, eliminated; with therapist training eliminated, sex was no longer significant. The dropout–continuer difference was nonsignificant for boys, but among girls, continuers improved significantly more than dropouts, $F(1, 23) = 7.08$, $p < .02$.

All other tests were nonsignificant in the 6- and 12-month samples: child age ignoring other factors, eliminating sex or therapist training; the Age × Sex and the Age × Therapy Type interactions; child sex ignoring other factors or with age or therapist training eliminated; the Sex × Therapy Type interaction, $F(2, 84) < 1$, $p > .50$, $F(2, 67) = 3.60$, $p < .05$, nonsignificant after Bonferroni adjustment; the main effect for therapy type ignoring other factors, $F(2, 87) = 1.51$, $p > .20$, $F(2, 70) = .38$, $p > .50$, or with age, sex, or therapist training eliminated; the main effect of therapist training with other factors ignored or with age or sex eliminated.

### *Overall Effect of Therapy on Outcome Measure 3: Teacher Reports*

We next analyzed TRF total problem *T* scores. Because of the small sample, we used only the Time 2 and 3 reports unadjusted for Time 1. At Time 1, dropouts and continuers did not differ significantly in the 6- or 12-month sample; thus, the strategy of not adjusting appeared acceptable. Dropouts and continuers did not differ significantly at Time 2 or 3. The ES was 0.31 at Time 2 and was 0.33 at Time 3. In a repeated-measures test, the effect of problem type was not significant at Time 2 or 3.

## Discussion

How effective is psychotherapy with children and adolescents? The present findings suggest a different conclusion than have meta-analyses of controlled-outcome studies (Casey & Berman, 1985; Weisz, Weiss, Alicke, & Klotz, 1987). On one measure, parents' ratings of their children's major problems, we did find one interaction indicating that girls who continued in therapy improved more than girls who dropped out. However, this interaction was found only in the 1-year follow-up and was nonsignificant when we controlled for therapist training. Otherwise, we found no evidence that children who completed a course of therapy improved any more than those who dropped out after intake. We also failed to replicate most of the predictive relations found in earlier meta-analyses: relations between child and adolescent outcomes and such factors as age, sex, problem type, therapy type, and therapist training, as well as interactions involving these factors.

Why such negative findings, findings so discrepant from results of meta-analyses? We have considered several possibilities. First, it is possible that the null findings resulted from excessive variability in the data. This seems unlikely for three reasons: (a) Reexamining our data with outliers removed had a negligible effect on our findings; (b) measures of the type used here have distinguished significantly between treated and untreated groups in controlled studies (e.g., Webster-Stratton, 1984); and (c) our data were sufficiently sensitive to reveal that both continuer and dropout groups showed significant improvement from Time 1 to 2 and Time 1 to 3 (all four *p*s < .01). Thus, null findings in the dropout–continuer comparisons resulted from the fact that both groups improved about equally, not from reduced power caused by excessive variability.

Second, it is possible that reliance on parent reports for two of our primary measures somehow biased findings in a null direction. Contrary to this possibility, the Weisz, Weiss, Alicke, and Klotz (1987) meta-analysis revealed no reliable differences between effect sizes generated by reports from parents, teachers, peers, and young clients themselves. In the Casey and Berman (1985) meta-analysis, parent ratings generated higher improvement scores than did teacher ratings (ES = 0.80 for parents and 0.19 for teachers); this would suggest that parent ratings might bias findings in the direction of apparent effectiveness of therapy, not apparent ineffectiveness. Finally, our teacher ratings provided a subsample of informants free of parental bias. The findings with this subsample gave little indication that therapy had positive effects.

A third possibility is that parents who opted to have their children drop out of therapy felt defensive and thus exaggerated how well their children were doing at Times 2 and 3. To explore this possibility, we focused on subjects for whom we had both parent (CBCL) and teacher (TRF) reports. Teachers should not be subject to such a hypothetical bias, so we tested whether parents of dropouts gave more positive reports on their children, in relation to teacher reports, than did parents of continuers. In an analysis of variance (ANOVA) of total problem *T* scores, we tested the Group (dropout vs. continuer) × Informant (parent vs. teacher) interaction at Times 2 and 3. Both results were nonsignificant; this casts doubt on the parental bias interpretation.

A fourth artifactual explanation is that dropouts (a) were actually in better condition psychologically than continuers at the outset of therapy or (b) for some other reason, improved after intake more than randomly selected control subjects would have improved. Arguing against Explanation (a) is our failure to find significant dropout–continuer differences on an array of demographic and clinical variables at intake. Arguing against Explanation (b) are the findings of recent analyses of the Weisz, Weiss, Alicke, and Klotz (1987) meta-analysis data: Weiss and Weisz (1988) found that randomly assigned control groups, too, improve significantly from pre- to posttreatment. In fact, the magnitude of improvement shown in the present sample, *z* (Time 1 to 2 divided by Time 1 *SD*) = 0.38, was very close to the average improvement shown in the random control groups of the meta-analysis ($z = 0.31$). Such findings suggest that our dropout group did not show inordinate improvement.

Is it possible, though, that dropouts did not continue treatment because their problems had been resolved or had grown less serious immediately after intake? It does not seem very

likely to us that so many youngsters who so closely resembled continuers demographically and clinically on the day of intake would have improved so sharply as to trigger clinic refusal within a matter of days and then slowed their improvement sufficiently that at 6 months and again at 1 year they could not be reliably distinguished from continuers on the outcome measures. But because dropouts were not randomly assigned, we cannot rule this possibility out. This area of uncertainty is one price paid for the enhanced external validity we gain by studying naturally occurring treated and untreated groups.

A fifth possibility is that our findings were influenced by the voluntary nature of subject participation. A number of parents did not participate at Times 2 and 3. One might argue that information from those parents would have changed our results. For example, parents of dropouts who improved may have been more likely to participate at Times 2 and 3 than parents of dropouts who did not improve, because the latter group might feel guilty. Such a pattern could make outcomes appear better than they actually were for the dropout group as a whole, thus masking true outcome differences between the full dropout and continuer groups. To assess the plausibility of a differential participation artifact, we calculated the CBCL changes that would have had to be true of children whose parents did not participate at Times 2 and 3, for the ES at Times 2 and 3 to equal the 0.79 ES reported in the meta-analysis by Weisz, Weiss, Alicke, and Klotz (1987). For the ES at Time 2 to equal 0.79, the ES for children with nonresponding parents would have had to be −1.63 for dropouts or +2.34 for continuers. For Time 3, the figures would have had to be −1.52 for dropouts or +1.61 for continuers. All four figures would place the average treated child of nonresponding parents above the 93rd percentile of the average untreated child of nonresponding parents, an unlikely occurrence. Thus, the artifactual process described here appears possible but not probable.

A final possibility must be considered: The findings presented here may be a valid reflection of the impact of child and adolescent psychotherapy as it typically occurs in outpatient clinics. If so, the findings would mean that therapy conducted by working clinicians, in actual clinics, with spontaneously referred children, under everyday conditions, may not be as effective as the therapy conducted under controlled conditions for research purposes. In fact, these findings give little evidence that such real-life therapy has any reliable effects. Certainly, the data do not clearly establish such a bleak conclusion. However, the findings do suggest that researchers and clinicians may need to reconsider the extent to which the findings of controlled-outcome studies with children can be generalized to therapy as it typically happens in clinics.

The present findings should not be read as evidence that meta-analytic findings are invalid. In fact, the positive findings of meta-analyses suggest that child and adolescent psychotherapy can be effective when conditions of therapy are carefully arranged, as when specific targets of treatment are clearly delineated, when these are well matched to the type of therapy provided, and when the therapists involved are well trained in the approach they use (see Beutler, 1979). Ultimately, controlled-treatment studies may help us in two significant ways: demonstrating that therapy can have positive effects and suggesting the conditions needed to produce those effects.

## References


Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile.* Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M., & Edelbrock, C. (1986). *Manual for the Teacher's Report Form.* Burlington, VT: Department of Psychiatry, University of Vermont.

Appelbaum, M. I., & Cramer, E. M. (1974). Some problems in the nonorthogonal analysis of variance. *Psychological Bulletin, 81,* 335–343.

Beutler, L. E. (1979). Toward specific psychological therapies for specific conditions. *Journal of Consulting and Clinical Psychology, 47,* 882–897.

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

Feild, H. S., & Armenakis, A. A. (1974). On the use of multiple tests of psychological significance in psychological research. *Psychological Reports, 35,* 427–431.

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

Kirk, R. E. (1982). *Experimental design: Procedures for the behavioral sciences* (2nd ed.). Monterey, CA: Brooks/Cole.

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

Levitt, E. E. (1971). Research on psychotherapy with children. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 474–494). New York: Wiley.

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

Webster-Stratton, C. (1984). Randomized trial of two parent-training programs for families with conduct-disordered children. *Journal of Consulting and Clinical Psychology, 52,* 666–678.

Weiss, B., & Weisz, J. R. (1988). *The effectiveness of psychotherapy with children and adolescents: A meta-analysis for researchers.* Unpublished manuscript, Vanderbilt University, Nashville, TN.

Weisz, J. R. (1986). Contingency and control beliefs as predictors of psychotherapy outcomes among children and adolescents. *Journal of Consulting and Clinical Psychology, 54,* 789–795.

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


Received September 27, 1988
Revision received April 24, 1989
Accepted June 2, 1989 ■