

Parent–Child Interaction Therapy: One- and Two-Year Follow-Up of Standard and Abbreviated Treatments for Oppositional Preschoolers

Reginald D. V. Nixon,^{2,4} Lynne Sweeney,² Deborah B. Erickson,³ and Stephen W. Touyz²

Received September 10, 2003; revision received December 10, 2003; accepted December 14, 2003

The long-term effect of two parent training programs for conduct problem preschoolers is reported. Families of 54 behaviorally disturbed preschool-aged children were randomly assigned to 1 of 3 treatment conditions: standard Parent–Child Interaction Therapy (STD), an abbreviated form of PCIT (ABB), and a no-treatment waitlist control group (WL). Of the families who completed treatment (STD and ABB), data were collected on 97% and 94% of families at 1- and 2-year follow-up, respectively. Follow-up assessment of parent report and independent observations indicated that treatment gains were largely maintained for both treatment conditions with little difference between the two treatments. The findings suggest that an abbreviated form of PCIT has long-term benefits for families with young children displaying early conduct problems.

KEY WORDS: preschoolers; ODD; externalizing behavior problems; treatment; PCIT.

Increasing attention is being paid to the preschool years for the early intervention of conduct problems (see Eyberg, 1992; Nixon, 2002; Webster-Stratton, 1993, for reviews). Parent–child interaction therapy (PCIT) is a promising intervention for externalizing problems in children (Foote, Eyberg, & Schuhmann, 1998), focusing on improving child–parent relationships and providing parents with skills to manage disruptive behavior. The short-term effectiveness of PCIT has been demonstrated in a number of controlled studies (see Nixon, 2002, for review), however long-term outcome is less established. For example, Eyberg et al. (2001) reported that at 2-year follow-up, 9 out of 13 families (69%) had sustained clinically significant changes. Funderburk et al. (1998), however, observed that behavioral gains which had generalized to the classroom setting immediately following treatment

had generally reverted to pretreatment levels by the 18-month follow-up.

Recently, Nixon, Sweeney, Erickson, and Touyz (2003) investigated the utility of abbreviating the standard format of PCIT to test whether this established treatment could be administered more easily. Thus oppositional preschoolers and their parents were randomly allocated to one of three groups: (a) standard PCIT (STD), (b) an abbreviated format of PCIT (ABB) which used instructional videotapes and telephone consultations during therapy, and (c) a no-treatment waitlist control group (WL). A group of nonproblem preschoolers and their families was recruited as a social validation comparison group (SV) as a further test of clinical significance. Both the STD and ABB therapies were superior to the WL condition immediately following treatment in terms of parent report of child behavior and independent assessment of parent–child interactions, and although there was a slight suggestion that the STD condition was superior to the ABB group immediately after treatment, these differences were negligible by 6-month follow-up. In this study, we were interested in whether the abbreviation of the standard format of PCIT maintained short-term gains in reducing

¹Portions of this research were presented at the Second Annual Parent–Child Interaction Therapy Conference, June, 2001, Sacramento, CA.

²Department of Psychology, University of Sydney, Sydney, Australia.

³Roberts Wesleyan College, Rochester, New York.

⁴Address all correspondence to Reginald D. V. Nixon, School of Psychology, Flinders University, GPO Box 2100, Adelaide, South Australia 5001, Australia; e-mail: reg.nixon@flinders.edu.au.

conduct-problem behavior in young children at 1- and 2-year follow-up.

METHOD

See Nixon et al. (2003) for full details of participants and methodology. Fifty-four clinical families completed treatment in either the STD condition ($n = 17$), ABB condition ($n = 20$), or WL condition ($n = 17$). Twenty-one families served in the SV condition. At the 1-year follow-up one family from the STD condition was not available for assessment. Two families were not available for the 2-year follow-up, one from the STD condition, the other from the ABB condition. Data are not reported for the WL condition because they were only assessed at pre- and posttreatment.

Participants

Local Child and Family Mental Health Teams, Early Childhood Centers, preschools, and newspapers were notified of a free treatment program for families who had preschool-aged children exhibiting behavioral difficulties. SV families were recruited by advertising in local papers. Inclusion and exclusion criteria were met through a telephone screen and assessment interview. Inclusion criteria for clinic families were: (a) the child was in the clinical range on the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999; Intensity score ≥ 132); (b) the child met diagnostic criteria for Oppositional Defiant Disorder (ODD) according to the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*; American Psychiatric Association, 1994); and (c) the primary referral problem was disruptive behavior which had been present for at least 6 months. Exclusion criteria were behavior problems due to organic pathology, trauma, history of severe physical or mental deficits, or receiving medication to manage behavioral difficulties. Children were accepted into the SV group if their parents reported that they were not experiencing difficulties in managing their child's behavior, the children's scores on the ECBI were in the normal range and they did not meet *DSM* criteria for ODD.

Demographic information on the intent-to-treat sample (ITT) at pretreatment was as follows. For the STD group, the children's mean age was 47.36 months ($SD = 6.26$) and consisted of 18 boys and 4 girls ($n = 22$). Mothers and fathers had a mean age of 34.73 ($SD = 4.54$) and 37.00 ($SD = 6.22$) years respectively. The ABB treatment consisted of 18 boys and 9 girls ($n = 27$), with a mean age of 48.30 months ($SD = 6.79$). Mothers and fathers in the ABB group had a mean age of 33.85 ($SD = 4.94$)

and 36.00 ($SD = 5.75$) years respectively. Across the two clinical groups, mothers and fathers had 12 years or less of education, and had an annual family income in the \$US23,000–40,599 range. There were four single parents each in the STD group and ABB group. Completers and noncompleters did not differ on any demographic or pretreatment variables. The SV group was comprised of 15 boys and 6 girls of mean age 44.71 months ($SD = 5.82$), with 19 children coming from partnered families. The mothers had a mean age of 35.52 ($SD = 5.00$) and fathers were aged 38.05 ($SD = 5.79$) on average. In terms of education, SV mothers and fathers had on average completed some technical/trade college, and their average income fell in the \$US23,200–40,599 range. All families spoke English at home and considered themselves "Australian" for demographic purposes. All but one family in the STD and SV group, and all but 3 families in the ABB group were Caucasian. The groups were equivalent on all demographic variables except for maternal education, with clinic mothers being less educated than SV mothers ($p < .05$).

Follow-Up Procedure

For the 1-year follow-up, questionnaires were completed by the parents, and a play observation of mother-child interaction was videotaped for independent coding. At 2-year follow-up, families were administered a diagnostic interview by phone by an experienced graduate-level research assistant unaware of group allocation, and questionnaires were returned by mail.

One-Year Follow-Up Measures

Unless otherwise specified, all measures are based on maternal report. Data were collected for fathers in partnered families using the ECBI. Mothers' and fathers' perceptions of children's oppositional behavior were measured using the Intensity Scale of the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999). The Child Domain of the Parenting Stress Index (PSI; Abidin, 1995) was used to measure stress associated with difficult qualities or characteristics of children that can lead to frustration and unfulfillment in the parenting role. Assessment of parenting styles was carried out using the Parenting Scale (PS; Arnold, O'Leary, Wolff, & Acker, 1993), a 30-item self-report scale designed to measure dysfunctional discipline practices in parents of young children. The Overreactivity factor was used in this study and assesses harsh, aggressive, and authoritarian discipline behaviors.

Independent observation of parent-child interactions in the clinic was conducted using the Dyadic

Parent-Interaction Coding Systems-II (DPICS-II; Eyberg, Bessmer, Newcomb, Edwards, & Robinson, 1994). There were five variables of interest: Total Number of Parental Praises, Total Number of Parental Criticisms, Total Number of Parental Commands, Total Child Compliance (reported as a percentage), and Total Child Deviant Behavior. All interactions were videotaped and randomized, then were coded by an undergraduate research assistant who was unaware of the aims of the study, group allocation and time of assessment. The research assistant received 40 hrs of training with the aid of the coding training manual recommended for DPICS-II (Eyberg, Edwards, Bessmer, & Litwins, 1994), and had to meet (or exceed) 80% agreement with the first author on two precoded tapes prior to coding the parent-child interactions. To maintain coding reliability and to guard against coder drift, weekly 1-hr meetings were held to review coding principles and to discuss any difficulties. The same coder was used in the Nixon et al. (2003) study.

Two-Year Follow-Up Measures

Mothers were interviewed using the parent form of the NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000) to determine the presence/absence of Oppositional Defiant Disorder according to DSM-IV. Diagnostic reliability was checked for 7 interviews (20%) by a second, independent rater, who was also blind to group allocation. There was 100% agreement ($\kappa = 1.0$) between the two raters.

Both parents (if applicable) completed the ECBI, and the Child Behavior Checklist (CBCL; Achenbach, 1991). The total Externalizing Scale (*T* score) was used for the CBCL. Given the association between parental depressive symptomatology and externalizing behavior problems (Lyons-Ruth, Easterbrooks, & Cibelli, 1997; Webster-Stratton & Hammond, 1988), parents also completed the Beck Depression Inventory-2 (BDI-2; Beck, Steer, & Brown, 1996). The total BDI-2 score was used in analyses.

Treatment

Both treatments (STD and ABB) were conducted as set out by the clinicians' guide for PCIT, *Parent Child Interaction Therapy* by Hembree-Kigin and McNeil (1995), with the recommended session components and durations closely followed. Treatment for both groups was administered individually. PCIT has two major phases. First, there is a focus on improving the child-parent relationship by teaching parents to play with their children in a

positive, nondirective manner, labelled "Special Playtime" (Session 1-5). During this phase parents are also taught basic skills (e.g., selective ignoring, labeled praise) to help modify unwanted behaviors of their child that would respond to such strategies. The focus of the second phase ("Discipline Skills," Session 6-12) was to teach parents skills and strategies in behavior management, using clear instructions in terms of commands to children and appropriate consequences for noncompliance (e.g., time-out and withdrawal of privileges). The STD treatment consisted of twelve 1-2 hr weekly sessions (depending on session content). For the ABB treatment, rather than a face-to-face training session to teach parents the skills of the two treatment phases, these skills were discussed and modeled on videotape by the first author (R.D.V.N) a copy of which was given to the family to watch at home. Five face-to-face sessions (1-2 hrs, depending on session content) were alternated with five 30-min telephone consultations to deliver intervention to the ABB group. Both the STD and ABB treatment included a 1-hr booster session (face-to-face) 1-month posttreatment. This booster consisted of checking progress, discussing remaining or new problem behaviors and strategies to handle such problems, and reviewing the rationale and skills of treatment components. The STD treatment took 15.5 hrs to administer, the ABB treatment, 9.5 hrs.

Treatment Integrity and Quality

Complete details of treatment integrity management can be found in Nixon et al. (2003). Briefly, therapy was conducted by the first author following the aims, objectives, and skills of each session as set out by Hembree-Kigin and McNeil (1995). Secondly, the therapist received ongoing supervision from the second author (L. S.). Sessions were videotaped (or recorded on audiocassette for phone sessions), and 20% were randomly selected and checked for treatment adherence by an undergraduate research assistant. Finally, two experienced clinical psychologists specialized in treating conduct problem families and working in a public treatment unit independent of the clinic where the study was conducted rated the quality/proficiency of 20% of the therapy sessions. As detailed in Nixon et al. (2003), these checks ensured both the fidelity and quality of treatment administered.

RESULTS

One-Year Follow-Up

Although clinic mothers were significantly less educated than SV mothers, maternal education was not

significantly correlated with any measures at either follow-up.

Long-term effects were examined through 2 (condition: STD, ABB) \times 2 (time: pretreatment, 1-year follow-up) repeated measures multivariate analysis of variance (MANOVA) for parent report and independent observations of mother-child interactions. (Because there were fewer fathers than mothers, and fathers only completed one measure (ECBI), inclusion of these data within the MANOVA would have reduced sample size and power. Thus fathers ECBI data were analyzed using a repeated measures analysis of variance [ANOVA]). Planned contrasts (t statistics) were used to compare follow-up scores between the treatment conditions. Two sets of analyses were conducted. First, intent-to-treat (ITT) analyses are reported for the entire sample, using the last observation carried forward (LOCF) method of replacement for missing data. For treatment nonstarters (families who completed the pretreatment assessment but did not attend therapy), dropouts (i.e., attended at least one therapy session), and families who completed treatment but failed to attend follow-up assessments, the conservative method of using pretreatment scores for LOCF was carried out. Second, to examine the effect of completing the treatments as prescribed, the same analyses were conducted with only those families who completed the 12-weeks therapies as scheduled. Statistically, there were negligible differences in outcome between the two analytic strategies. Table I summarizes the means and standard deviations and effect sizes for relevant dependent variables at each assessment. Hedges's g effect sizes (Hedges, 1982) were calculated, then adjusted to unbiased Hedges's g to correct for the small sample size (Hedges, 1982; Rosenthal, 1991).

For the ITT sample, there was a significant main effect of time for mothers' report of child behavior (ECBI, PSI, PS), $F(3, 42) = 15.37, p < .001$ and for independent observations (mother praise, mother criticisms, mother commands, child compliance, child deviant behavior), $F(5, 39) = 7.25, p < .001$. Main effects for group or interaction of time and group were not significant, suggesting the STD and ABB treatment groups showed a similar pattern of change. The same pattern of findings was observed for the completers sample, with only a significant effect of time being observed on mothers' report of child behavior, $F(3, 31) = 23.83, p < .001$ but no group or interaction effects. Similarly, fathers' report of child behavior (ECBI) showed a significant effect of time for both the ITT and completers sample, $F(1, 36) = 25.07, p < .001$ and $F(1, 27) = 32.95, p < .001$, respectively, but nonsignificant group and interaction effects.

Paired t tests between pretreatment and 1-year follow-up for the ITT sample generally showed that for both treatments, child behavior was improved at follow-up compared with baseline levels (see Table I for significance details). Some exceptions were noted. In the STD group, there was not a significant decrease in child deviant behavior according to independent observations. In the ABB group at follow-up, mothers' critical statements, and children's compliance and deviant behavior were not significantly different from their baseline levels. Independent t tests revealed only one significant difference between the groups at 1-year follow-up, with ABB mothers observed to praise their children significantly more than STD mothers, $t(43) = 2.30, p < .05$. When the completers sample was examined in the same fashion, all but one of the above findings was replicated. The exception was that there was not a significant difference between the STD and ABB group in terms of maternal praise at 1-year follow-up.

Clinical Significance at One-Year Follow-Up

The same indices were used as in the original report: mothers' report of conduct problems and parenting stress (ECBI, PSI), and three independent observations of mother-child interactions: maternal commands and criticisms, and child compliance. Children were classified as having made clinically significant improvement if mothers' report of behavior or stress fell below the published clinical cut offs for that measure (132 for the ECBI, Eyberg & Pincus, 1999, and 116 for the PSI, Abidin, 1995) and if this change was determined to be reliable as set out by Jacobson and Truax (1991). Jacobson and Truax define a reliable change by a change of more than 1.96 *SEMs* between pretreatment and posttreatment (or follow-up). A 30% reduction in commands or criticisms and 30% increase in child compliance at posttreatment (compared with baseline level) was also used as an index of change for mother behaviors. The two treatments were compared with each other, then with the SV group on the five variables of interest. Only children who were in the clinical range on the PSI at pretreatment were considered in analyses for this measure. (All children were in the clinical range on the ECBI at pretreatment).

Analyses indicated that there were no significant differences in the ITT sample between the two treatments at 1-year follow-up. Thus 50% of STD children were reported by their mothers to have made reliable and clinically significant changes in oppositional behavior (ECBI) compared with 42% of ABB children, $\chi^2(1, N = 48) = .28, ns$. Forty-five percent of mothers in the STD condition reported clinically significant reductions in parenting

Table 1. Means, Standard Deviations, Dependent *t* Tests, and Effect Sizes for Pretreatment and 1-Year Follow-Up Assessment

Parent report	Measures																		STD vs. ABB <i>g</i>				
	STD						ABB						SV							Pre vs. FU			
	Pre			F/U			Pre			F/U			Pre			F/U				STD		ABB	
	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>			<i>t</i>	<i>df</i>	<i>t</i>	<i>df</i>
ECBI (M)	167.50	20.01	133.41	36.75	158.92	20.30	136.42	29.52	107.56	15.94	110.89	24.78							-4.65***	21	-4.43***	23	0.09
Completer	166.19	19.48	119.31	30.65	155.84	17.15	127.42	22.78											-5.89***	15	-4.99***	18	0.30
ECBI (F)	150.58	28.34	122.16	39.63	141.00	24.31	122.53	31.79											-3.68**	18	-3.48**	18	0.01
Completer	151.43	22.22	112.86	34.73	137.20	22.71	113.80	27.22											-4.27**	13	-3.82**	14	0.03
PSI	133.82	16.51	114.45	23.21	135.79	24.44	119.17	26.35	88.65	12.37	90.72	21.59							-4.47***	21	-5.12***	23	0.19
Completer	136.31	10.83	109.69	20.62	132.95	21.77	111.95	19.35											-5.52***	15	-6.06***	18	0.11
PS	3.48	0.65	3.05	0.95	3.63	0.80	3.24	1.07											-2.66*	21	-3.42**	23	0.18
Completer	3.58	0.64	2.99	1.05	3.70	0.73	3.20	1.10											-2.81*	15	-3.66**	18	0.19
Mother-child behaviors																							
Praises	7.33	6.51	12.62	7.87	12.75	12.74	21.04	15.04											3.76**	20	3.44**	23	-0.68
Completer	8.87	6.96	16.27	5.95	10.94	12.75	22.65	16.04											4.40**	14	3.83**	16	-0.28
Criticisms	6.38	6.40	3.67	4.43	6.46	5.79	5.08	4.96	2.73	3.56	3.00	2.62							-2.54*	20	-1.20	23	0.29
Completer	6.60	6.56	2.80	3.14	5.65	5.22	3.71	3.14											-2.69*	14	-1.20	16	0.50
Commands	40.00	21.03	24.67	13.01	44.50	22.25	32.96	15.93	26.33	11.70	27.53	10.44							-3.55**	20	-3.21**	23	0.56
Completer	43.20	22.13	21.73	10.29	45.06	24.68	28.76	14.13											-4.06**	14	-3.53**	16	0.55
Child compliance	0.60	0.27	0.73	0.28	0.58	0.23	0.67	0.25	0.68	0.23	0.70	0.21							3.51**	20	1.72	23	0.22
Completer	0.65	0.21	0.83	0.16	0.58	0.26	0.70	0.29											4.00**	14	1.75	16	0.55
Child deviance	12.14	20.86	5.52	6.5	12.04	16.87	9.08	13.60											-1.62	20	-1.20	23	0.32
Completer	13.73	23.90	4.47	4.03	10.00	14.21	5.82	5.91											-1.65	14	-1.20	16	0.26

Note. STD = standard treatment; ABB = abbreviated treatment; SV = social validation Group; Pre = pretreatment, F/U = 1-year follow-up; ECBI (M) = Eyberg Child Behavior Inventory (Intensity score) for mother; ECBI (F) = Eyberg Child Behavior Inventory (Intensity score) for father; PSI = Parenting Stress Index (child domain score); PS = Parenting Scale (overreactivity score); Criticisms during DPICS-II by mother (total no.); Praises during DPICS-II by mother (total no.); Commands during DPICS-II by mother (total no.); Child compliance during DPICS-II (percentage of commands obeyed); Child deviance during DPICS-II (total no. of deviant acts); *g* = Hedges's unbiased *g* effect size. A positive effect size indicates better outcome for the STD condition on that measure relative to the ABB condition.

* $p < .05$. ** $p < .01$. *** $p < .001$.

stress (PSI) compared with 25% of ABB mothers, $\chi^2(1, N = 40) = 1.76, ns$. In terms of independent observations, 48 and 38% of mothers in the STD and ABB groups made clinically significant reductions in the number of commands they gave their children, and 57 and 58% of mothers reduced the number of criticisms, $\chi^2(1, N = 45) = .47, ns$, and $\chi^2(1, N = 45) = .01, ns$, respectively. Fourteen percent of children in the STD condition and 21% in the ABB condition were observed to have made clinically significant improvements in compliance at the 1-year follow-up, $\chi^2(1, N = 45) = .33, ns$.

Although no differences were observed between the two treatments when the completer sample was tested for clinically significant changes, not surprisingly, greater proportions of change were noted than in the ITT sample. Thus the following proportions of families made clinically significant changes in the completer sample: on the ECBI, 69% in the STD condition versus 58% in the ABB condition, $\chi^2(1, N = 35) = .44, ns$; on the PSI, 56% in the STD condition versus 31% in the ABB condition, $\chi^2(1, N = 32) = 2.03, ns$; maternal commands, 67% versus 53%, $\chi^2(1, N = 32) = .62, ns$; maternal criticisms, 60% versus 59%, $\chi^2(1, N = 32) = .01, ns$; and child compliance, 20% versus 29%, $\chi^2(1, N = 32) = .38, ns$.

When comparisons were made between the ITT sample and SV group, mothers from ABB group continued to report behavior (ECBI) at a level significantly higher than that of SV mothers, $t(40) = 2.97, p < .01$. The same finding was observed for both treatments for parental stress (PSI); for the STD condition, $t(38) = 3.32, p < .01$, for the ABB condition, $t(38) = 3.73, p < .01$. The number of commands and criticisms given by STD and ABB mothers, however, was not significantly different from SV mothers, and children in the STD and ABB groups were not significantly different from SV children in terms of their compliance. The same pattern of results was demonstrated when data for the completer sample was analyzed.

Equivalency Testing at 1-Year Follow-Up

The approach of Rogers, Howard, and Vessey (1993) was adopted to test the *similarity* of the two treatments. In such testing, a predetermined nonzero difference value (also known as an "equivalence interval") is defined as being a difference between two treatments that is considered to be clinically unimportant. The researcher then tests the null hypothesis that the difference between the two treatment means on a particular variable is equal or larger than the specified difference. Equivalency is derived statistically when the null hypothesis is rejected and the alternative hypothesis, that the difference between the two groups is smaller than the predetermined difference,

is accepted. An interval of 1 *SD* (pooled) for the relevant variable was used.

For the ITT sample, equivalency was established for the following measures: mother ECBI ($z = 3.45, p < .01$); father ECBI ($z = 3.42, p < .01$); PSI ($z = 2.54, p < .01$); PS ($z = 2.63, p < .01$); maternal criticisms ($z = 3.48, p < .01$); and child compliance ($z = 2.78, p < .01$). A significant difference was established between the two groups in terms of maternal praise, with ABB mothers praising significantly more than STD mothers at 1-year follow-up. Equivalency was not observed between the two treatments in terms of maternal commands and child deviance. Rogers et al. (1993) state that when both traditional and equivalence tests fail (i.e., the null hypothesis is retained for both tests as in the case on the latter two variables), insufficient evidence exists to make a decision about equivalency for reasons such as excessive variance or low power. For the completer sample, mother and father report of behavior (ECBI; $z = 2.49, p < .01$, and $z = 2.93, p < .01$), mother report of parenting stress (PSI; $z = 2.71, p < .01$) and discipline (PS; $z = 2.30, p < .05$) was equivalent between the STD and ABB treatments. Similarly, equivalency was established for independent assessment of maternal criticisms, $z = 2.00, p < .05$. Equivalency was not observed between the two treatments, however, for number of maternal commands, child compliance, or child deviance.

Two-Year Follow-Up

All but one family in each group of the treatment completers were available for follow-up, (STD = 16, ABB = 19 families). Two families in the STD group and three in the ABB group had sought additional assistance for their child's behavior problems since the 1-year follow-up. Because all but one of the 2-year follow-up measures had not been previously administered (the ECBI being the exception), the two groups were compared at the single time point (i.e., at 2-year follow-up) with *t* tests and chi-square analyses. All mothers were interviewed with the DISC, providing diagnostic information regarding their children. The return rate for questionnaires (ECBI, CBCL, BDI-2), however was modest and the results are interpreted cautiously. Only 10 of the 16 mothers and fathers in the STD group returned questionnaires, and nine and six mothers and fathers respectively in the ABB group. Participation rate of the SV group at 2-year follow-up was similarly poor, with only five families returning questionnaires. SV families were therefore not considered in the 2-year follow-up analyses. Families who did return questionnaires were compared to those who did not on demographic, pretreatment, 1-year follow-up variables,

and child diagnosis at 2-year follow-up. Only one difference was observed, with fathers who returned questionnaires being significantly older than the nonresponding fathers ($p < .05$). Fathers' age, however, was not correlated with any outcome variable and thus was not considered further.

Independent samples t tests revealed no significant differences between the STD and ABB condition at follow-up (see Table II for details). Mean scores were all in the normal range. Examination of unbiased Hedges's effect sizes for STD-ABB treatment comparisons indicates that three of the four measures assessing child behavior problems appeared to favor the ABB group. By Cohen's standards (Cohen, 1988), these effect sizes range from small to medium. At 2-year follow-up, 56% of STD children and 68% of ABB children still met criteria for Oppositional Defiant Disorder (assessed using the DISC), but this difference was not significant, $\chi^2(1, N = 35) = .55, ns$. As we had pretreatment scores for the ECBI, we were able to examine the clinical significance of change using this measure. In apparent contrast to the proportion of children meeting criteria for ODD at follow-up, 70 and 67% of STD and ABB children were reported by their mothers to have made reliable and clinically significant reductions in oppositional behavior, $\chi^2(1, N = 19) = .02, ns$.

In terms of equivalency testing, father report of behavior (CBCL; $z = 1.67, p < .05$) was equivalent between the STD and ABB treatments. Although equivalency was not observed between the two treatments on the remaining measures, lack of power appeared to be an issue, as three of the remaining five tests came close to signif-

icance (mother ECBI; $z = 1.57, p = .058$; father ECBI; $z = 1.53, p = .063$; father BDI-2; $z = 1.63, p = .052$).

DISCUSSION

The results generally indicated that both the STD and ABB treatments had maintained their short-term gains at both 1- and 2-year follow-up. The current findings are also in accord with our original report which indicated that by 6-month follow-up, few differences were apparent between the two treatment approaches. By using traditional significance testing, equivalency testing, and examining the clinical significance of change, we can be reasonably confident that these results are robust. At the 1-year follow-up, both independent observations and parent-report data suggested that children from both treatments were causing their mothers less stress, were less oppositional, and that their mothers had maintained their reduction of coercive discipline techniques. Although no differences were evident between the two treatments in terms of the percentage of families making clinically significant changes, greater proportions were observed in the STD condition. It is also worth noting that independent observational data demonstrated that the STD group was significantly improved on four of the five variables at 1-year follow-up compared with their baseline levels; the ABB group was improved only on two. We highlight the fact that somewhat small proportions of children appeared to have maintained a 30% increase in compliance compared with baseline levels (e.g., 20% of STD children and 29% of ABB children in the completer sample). This finding deserves explanation. First, children had relatively good rates of compliance at baseline (around 60% compared with the 68% compliance rate observed in SV children), therefore had a ceiling in terms of improving by 30% or more. Second, rates of compliance at 1-year follow-up in both the clinical groups were comparable to nonproblem children in the SV group suggesting that treatment children were similar to their nonproblem peers in terms of compliance.

Although limited to treatment completers, and accompanied by the *caveat* of a reduced response rate for questionnaire measures, the 2-year follow-up provided some intriguing results. First, although there were no differences statistically or clinically between the two treatments, effect sizes for parent report measures of child behavior appeared to favor the ABB condition in contrast to the effect sizes observed at the 1-year follow-up. Second, we observed an interesting discrepancy between clinically significant changes being observed for approximately 70% of the children on the ECBI, yet around 60% of children appearing to meet criteria for ODD on the

Table II. Means, Standard Deviations, Independent t Tests, and Effect Sizes for Completer Sample at 2-Year Follow-Up

	Measures						
	STD		ABB		STD vs. ABB		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>g</i>
<i>Mother report</i>							
ECBI	126.50	23.39	120.44	24.75	0.55	17	-0.24
CBCL	54.70	11.42	62.11	8.58	-1.58	17	0.44
BDI-2	6.30	3.34	10.33	10.49	-1.16	17	0.51
<i>Father report</i>							
ECBI	117.50	33.73	110.17	32.87	0.43	14	-0.21
CBCL	55.10	10.54	52.33	9.69	0.52	14	-0.26
BDI-2	5.70	5.23	4.33	3.14	0.58	14	-0.28

Note. STD = standard treatment; ABB = Abbreviated Treatment; ECBI = Eyberg Child Behavior Inventory (intensity score); CBCL = Child Behavior Checklist (externalizing t score); BDI-2 = Beck Depression Inventory-2 (total score). g = Hedges's unbiased g effect size. A positive effect size indicates better outcome for the STD condition on that measure relative to the ABB condition.

DISC. One explanation for this discrepancy is that repeated administration of the ECBI resulted in some form of demand characteristics or positive expectancy of therapy. Although we acknowledge this as a valid hypothesis, it could be argued that a high demand for positive reporting would possibly be more likely to occur in an interview context (DISC) rather than with a parent completing a questionnaire at home and returning it via mail (as was done with the ECBI). We recognize that without other comparison measures, it is difficult to determine what best accounts for this discrepancy. It should be remembered, however, that the ECBI measures problematic behavior on a continuum (as opposed to the dichotomy of a diagnosis) and was analyzed in a more sophisticated (and perhaps more meaningful) manner to determine whether clinically significant reductions of oppositional behavior had occurred. When we compared scores on the questionnaires used in both the present study and that of Eyberg et al. (2001), we observed that both studies had comparable means at follow-up on the ECBI, CBCL, and PSI (although pretreatment scores were higher in the Eyberg sample). Similarly, Hood and Eyberg (2003) observed that 70% of their sample demonstrated clinically significant changes on the ECBI at 3–6 year follow-up, and reported mean ECBI scores consistent with ours. Coming from an independent research unit, the present findings provide further support of PCIT as an effective program for young conduct-problem children.

Overall, the present findings of the long-term effectiveness of PCIT are comparable to those reported by Eyberg et al. (2001) and Hood and Eyberg (2003). In addition, there is also some consistency of results between the current study and previous research where the maintenance of effects of other parent training programs for young conduct-problem children has been evaluated two to three years posttreatment (e.g., Reid, Webster-Stratton, & Hammond, 2003; Webster-Stratton, 1990). We concede, however, limitations of the study. First, ethical considerations led to the treatment of the waiting list control group. Without such a comparison group at follow-up, it cannot definitely be concluded that improvements in behavior in the treated groups were solely because of the effects of treatment and not maturational factors or instability of preschool behavior problems. Second, teacher reports of child behavior were not obtained which would have provided a valuable source of information independent of parent report. Finally, the study's modest sample size must be acknowledged. Clearly, replication with increased numbers is critical to properly examine the clinical relevance of an abbreviated PCIT.

In summary, despite the positive finding that a relatively brief intervention can maintain its effects up to 2

years later, it is clear that not an insignificant number of children and their families continue to experience substantial difficulties. Although somewhat of a truism, this study highlights the importance of researching why some families do not respond to treatment, and then translating these findings into interventions that specifically address these issues.

ACKNOWLEDGMENTS

The authors thank all the families who participated in the study and the following people who assisted with the research (in alphabetical order): Marita Brack, Margaret Charles, Rose Cumberland, Karyn Levy, and Gerri Minshall.

REFERENCES

- Abidin, R. R. (1995). *Parenting Stress Index—Professional manual* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment, 5*, 137–144.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory manual* (2nd ed.). San Antonio, TX: Psychological Corporation.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Eyberg, S. M. (1992). Assessing therapy outcome with preschool children: Progress and problems. *Journal of Clinical Child Psychology, 21*, 306–311.
- Eyberg, S. M., Bessmer, J., Newcomb, K., Edwards, D., & Robinson, E. A. (1994). *Dyadic parent-child interaction coding system-II: A manual* (Social and behavioral sciences documents Ms. No. 2897). San Rafael, CA: Select Press.
- Eyberg, S. M., Edwards, D., Bessmer, J., & Litwins, N. (1994). *The workbook: A coder training manual for the dyadic parent-child interaction coding system-II* (Social and behavioral sciences documents Ms. No. 2898). San Rafael, CA: Select Press.
- Eyberg, S. M., Funderburk, B. W., Hembree-Kigin, T. L., McNeil, C. B., Querido, J. G., & Hood, K. K. (2001). Parent-child interaction therapy with behavior problem children: One and two year maintenance of treatment effects in the family. *Child and Family Behavior Therapy, 23*, 1–20.
- Eyberg, S. M., & Pincus, D. (1999). *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory—Revised: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Foote, R., Eyberg, S., & Schuhmann, E. (1998). Parent-child interaction approaches to the treatment of child behavior problems. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 20, pp. 125–151). New York: Plenum.
- Funderburk, B. W., Eyberg, S. M., Newcomb, K., McNeil, C. B., Hembree-Kigin, T., & Capage, L. (1998). Parent-child interaction therapy with behavior problem children: Maintenance of treatment effects in the school setting. *Child and Family Behavior Therapy, 20*, 17–38.

- Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, *92*, 490–499.
- Hembree-Kigin, T. L., & McNeil, C. B. (1995). *Parent-child interaction therapy*. New York: Plenum.
- Hood, K. K., & Eyberg, S. M. (2003). Outcomes of parent-child interaction therapy: Mothers' reports of maintenance three to six years after treatment. *Journal of Clinical Child and Adolescent Psychology*, *32*, 419–429.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.
- Lyons-Ruth, K., Easterbrooks, M. A., & Cibelli, C. D. (1997). Infant attachment strategies, infant mental lag, and maternal depressive symptoms: Predictors of internalizing and externalizing problems at age 7. *Developmental Psychology*, *33*, 681–692.
- Nixon, R. D. V. (2002). Treatment of behavior problems in preschoolers: A review. *Clinical Psychology Review*, *22*, 525–546.
- Nixon, R. D. V., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2003). Parent-interaction therapy: A comparison of standard and abbreviated treatments for oppositional preschoolers. *Journal of Consulting and Clinical Psychology*, *71*, 251–260.
- Reid, M. J., Webster-Stratton, C., & Hammond, M. (2003). Follow-up of children who received the Incredible Years intervention for oppositional-defiant disorder: Maintenance and prediction of 2-year outcome. *Behavior Therapy*, *34*, 471–491.
- Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin*, *113*, 553–565.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
- Shaffer, D., Fisher, P. S., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 28–38.
- Webster-Stratton, C. (1990). Long-term follow-up of families with young conduct problem children: From preschool to grade school. *Journal of Clinical Child Psychology*, *19*, 144–149.
- Webster-Stratton, C. (1993). Strategies for helping early school-aged children with oppositional defiant and conduct disorders: The importance of home-school partnerships. *School Psychology Review*, *22*, 437–457.
- Webster-Stratton, C., & Hammond, M. (1988). Maternal depression and its relationship to life stress, perceptions of child behavior problems, parenting behaviors, and child conduct problems. *Journal of Abnormal Child Psychology*, *16*, 299–315.