

① - article
Comments

EVIDENCE BASED PRACTICE

Adams (2001)

Evaluating the Quality of Evidence

Critically Appraising Articles About Treatment

1. Determine relevance based on conclusion of abstract.

Is the article worth taking the time to read?

1.1. Did the authors study an outcome that clients would care about?

Yes (go on)

No (stop)

1.2. Is the problem studied one that is common to your practice?

Yes (go on)

No (stop)

1.3. Is the intervention feasible?

Yes (go on)

No (stop)

1.4. Will this information, if true, require you to change your current practice?

Yes (go on)

No (stop)

Perhaps, use the program to train or
Shape current programs?

2. Determine the validity of a single study.

If the answers to all four of the preceding questions are "Yes", then continued assessment of the article is mandatory. Study design flaws are common; fatal flaws are arresting.

2.1. Was the assignment of clients to treatment randomized?

yes

2.2. If random assignment was ensured, was the randomization list concealed from those assigning clients to treatment from those evaluating outcomes?

don't know, doesn't

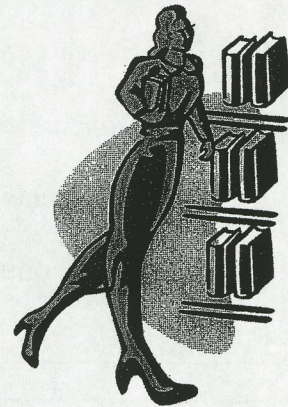
2.3. Were all subjects who entered treatment accounted for in its conclusion?

say

Some may be lost to follow-up when outcome is measured, but this is not fatal if a sufficient number were accounted for.

yes - though some did
drop out

1



EVIDENCE BASED PRACTICE

Evaluating the Quality of Evidence

Critically Appraising Articles About Treatment

2.4. Were subjects analyzed in the groups to which they were randomly assigned?

Yes

2.5. Were subjects and treatment personnel "blind" to which treatment was being received?

don't know - doesn't say

2.6. Aside from the experimental treatment, were the groups treated equally?

Yes - as far as we can tell

2.7. Were the groups similar at the start of the trial?

Yes

3. Determine the impact of the treatment to judge whether the treatment is worth the effort.

		Outcome Event		Total	Attrition!
		Yes	No		
Control (Not treated or given alternative treatment)	a=20 b=3 c=13 d=8	a 3	b 20	a+b 23 23	Improvement in the General Functioning Scale is the outcome of interest
Experimental Group (Given treatment)		c 8	d 13	c+d 21	

Flip Flop

35 is total number in the study

3.1. Control event rate (CER) or the risk of an outcome event in the control group.

$$CER = a/(a + b)$$

$$3/(23) = .13$$

$$20/23 = .87$$

3.2. Experimental event rate (EER) or the risk of an outcome event in the experimental group.

$$EER = c/(c + d)$$

$$8/(21) = .38$$

$$13/21 = .62$$

these are right!

3.3. Relative risk reduction (RRR)

$$RRR = (CER - EER)/CER$$

$$(.13 - .38) / .13 = -1.92$$

$$(.87 - .62) / .87 = .29$$

No negative number.

3.4. Absolute risk reduction (ARR)

$$ARR = CER - EER$$

$$= .87 - .62 = .25$$

3.5. Number needed to treat (NNT) or number treated who will not benefit from treatment in order to treat one individual who will benefit.

$$NNT = 1/ARR$$

$$1/.25 = 4$$

EVIDENCE BASED PRACTICE

Evaluating the Quality of Evidence

Critically Appraising Articles About Treatment

- 3.6. 95% Confidence interval (CI) for NNT or the estimated range of values for NNT with 95% chance that the true value is within this range.

$$95\% \text{ CI} = \pm 1.96 \sqrt{\frac{\text{CER} \times (1 - \text{CER})}{\# \text{ of control subjects}} + \frac{\text{EER} \times (1 - \text{EER})}{\# \text{ of exper. subjects}}}$$

$$\sqrt{\frac{.87 \times (1 - .87)}{23} + \frac{.62 \times (1 - .62)}{21}}$$

$$\sqrt{.0049 + .0112}$$
$$\pm 1.96 (.1269)$$
$$4 \pm .2487$$

4. Should I apply these valid, important results to my client or clients?

- 4.1. Do these results apply to my client(s)?

- 4.1.1. Is my client so different from those in the trial that the results do not apply?

No mention of Race! This is most married females (parents) with H.S. diploma

- 4.1.2. Is the benefit to my client(s) worth the cost, risk, and effort?

Perhaps, looks promising

- 4.2. Are my client(s) values and preferences satisfied by the interventions offered

AS far as we can tell - Maybe. Need more info.

- 4.2.1. Do I have a clear assessment of the client(s) values and preferences?

Good question! - Do we?

- 4.2.2. Does this intervention and its potential consequences serve these values?

Discussion needed.

△ - article
Comments

Adams (2001)

EVIDENCE BASED PRACTICE
Evaluating the Quality of Evidence
Quality of Study Rating Form (QSRF)

Client type(s): Parents - mostly moms, No info. on race/ethnicity!

Intervention method(s): Parent training + mental hltH srvcS vs M.h. srvcS alone

Outcome measure to compute ES1: Overall general functioning (FAD) *could use whole scale or part*

Outcome measure to compute ES2: affective responsiveness (FAD) *(OV)*

Outcome measure to compute ES3: behavior control (FAD) *(OV)*

Source in APA format: Adams, J.F. (2001) Impact of fami Parent training on family functioning. Child & Family Behavior Therapy, 23(1) 29-42

Criteria for Rating Study

Clear Definition of Treatment					6. Subjects randomly assigned to treatment or control (10 pts.)	7. Analysis shows equal treatment and control groups before treatment (5 pts.)	8. Subjects blind to being in treatment or control group (5 pts.)
1. Who (4 pts.)	2. What (4 pts.)	3. Where (4 pts.)	4. When (4 pts.)	5. Why (4 pts.)			
3	4	1	4	4	10	3	2

↳ No race/ethnicity

↳ Not much detail

↳ doesn't really say

Criteria for Rating Study (cont.)

9. Subjects randomly selected for study inclusion (4 pts.)	10. Control or non-treated group used (4 pts.)	11. Number of subjects in smallest treatment group exceeds 20 (4 pts.)	12. Outcome measure has face validity (4 pts.)	13. Treatment outcome measure was checked for reliability (5 pts.)	14. Reliability measure has value greater than .70 or percent of rater agreement greater than 70% (5 pts.)
4	4	4	4	5	4

↳ says the reliability is "adequate"

Criteria for Rating Study (cont.)

15. Those rating outcome rated it blind (10 pts.)	16. Treatment outcome was measured after treatment was completed (4 pts.)	17. Test of statistical significance was made and $p < .05$ (10 pts.)	18. Follow-up was greater than 75% (10 pts.)	19. Total quality (add 1-18)	20. Effect size = (ES1) = SD units = (mean of treatment - mean of control or alternate treatment) ÷ (standard deviation of control or alternative treatment)
0	4	10	6	76	$\frac{2.06 - 1.98}{.59} = .136$

↳ does not say

↳ a lot of attrition for a small sample!

Criteria for Rating Effect Size

21. Effect size (ES2) = Absolute risk reduction = (percent improved in treatment) - (percent improved in control)	22. Effect size (ES3) = Number needed to treat = $100 \div ES2$
$\frac{4.00 - 1.98}{.68} = 38\% - 12\% = 26\%$	$100 / 26 = 3.85$

Calculating effect size depends on whether you are trying to increase or decrease an event!

EVIDENCE BASED PRACTICE

Evaluating the Quality of Evidence

Quality of Study Rating Form (QSRF)

The QSRF:

- Takes the guesswork out of rating effectiveness studies.
- Is designed so best effectiveness studies score highest.
- Includes indices of study quality and treatment impact.
 - Total quality points – provides an index of confidence in the studies validity.
 - Indices that may be compared across studies to estimate the relative magnitude of a treatment's effect.
 - Standardized mean difference (ES1)
 - Absolute risk reduction (ARR or ES2)
 - Number needed to treat (NNT or ES3)
- Instructions for scoring.
 - Items 1 to 18 assess quality. These are summed in item 19. Item 19 ranges from 0 to 100. The closer to 100, the more confidence the rater can place in the study's findings.
 - Items 20, 21, and 22 are three relative indices of treatment effect size. These indices summarize the impact of treatment in standardized units.
- Explanation of criteria regarding study quality.
 - The first section of the QSRF states the identifying information regarding the client type (e.g., depressed, middle-aged men); intervention method (e.g., aerobic exercise four or more times a week); and the outcome measure(s) – preferably ones that can be used to calculate treatment effect size (e.g., score on Beck Depression Inventory).
 - Experience with the QSRF's reliability indicates that reliability will be highest for all-or-nothing points for each item; so give either zero points or the particular point value indicated if the study meets the criterion, as numbered on the form and described in the following list:
 1. *Who*: The authors describes who is treated by stating the subject(s)' average age *and* standard deviation of age, *and* sex or proportion of males and females, *and* clearly defines clients' presenting problem(s).
 2. *What*: The authors tell what the treatment involves so specifically that you could apply the treatment with nothing more to go on than their description, *or* they refer you to a book, videotape, CD-ROM, article, or Web address that describes the treatment method.
 3. *Where*: Authors state where the treatment occurred so specifically that you could contact people at the facility by phone, letter, or E-mail address.
 4. *When*: Authors tell the *when* of treatment by stating how long subjects participated in the treatment in days, weeks, or months *or* tell how many treatment sessions were attended by subjects.
 5. *Why*: Authors either discuss a specific theory that describes why they used one or more treatment methods, *or* they cite literature that supports the use of the treatment method.
 6. *Subjects randomly assigned to treatment or control*: The author states specifically that subjects were *randomly assigned* to treatment groups or refers to the assignment of subjects on the basis of random numbers, computer algorithm, or accepted randomization procedures. This means that the procedure resulted in subject having an equal chance of being assigned to treatment or control groups.
 7. *Analysis shows equal treatment and control groups before treatment*. Even though subjects have been randomly assigned, unequal treatment and control groups can occur by chance; so, to guard against this, the authors need to make comparisons across treatment and control groups on key client characteristics to see that they are similar prior to treatment (e.g., sex, race, age, economic status, condition, strengths).
 8. *Subjects blind to being in treatment or control group*: Subjects who know they are in a control group can experience effects of being there including demoralization or competition with experimentals. Subjects who know they are in a treatment group can experience powerful healing effects because they expect them. Give points for subjects blinded if two or more groups get some kind of treatment, if controls get

EVIDENCE BASED PRACTICE
Evaluating the Quality of Evidence
Quality of Study Rating Form (QSRF)

some form of sham treatment that is not expected to have an effect but gives assurance to subjects that something is being done, if subjects serve in a delayed treatment control group where they serve as controls but get treatment later, or if subjects truly do not know whether they are in a treatment or control group.

9. *Subjects randomly selected for inclusion in study:* Selection of subjects is different from *random assignment*. *Random selection* means that subjects are taken from some potential pool of subjects for inclusion in the study by using a table of random numbers or other statistically random procedures. For example, if subjects are chosen randomly from among all residents on a psychiatric ward, the results of the study can be generalized more confidently to all residents of that ward.
10. *Control (nontreated) group used:* Member of a *nontreated control group* do not receive a different kind of treatment; they receive *no* treatment. An example of a non treated control group would be a group of subjects who are denied counseling while others are given group counseling. Subjects in nontreated control group may receive treatment at a later date but do not receive treatment while experimental group subjects are receiving their treatment.
11. *Number of subjects in smallest treatment group exceeds 20:* Those in the treatment group or groups are those who receive some kind of special care intended to help them. It is this treatment that is being evaluated by those doing the study. In order to meet criterion 11, *the number of subjects in the smallest treatment group must be at least 21*. Here, *number of subjects* means total number of individuals, not number of couples or number of groups.
12. *Outcome measure has face validity:* Face validity is present if the outcome measure used to determine the effectiveness of treatment makes sense to you. A good criterion for the sense of an outcome measure is whether the measure evaluates something that should logically be affected by the treatment. For example, drinking behavior has face validity as an outcome measure for treating alcoholism.
13. *Treatment outcome measure was checked for reliability:* For this criterion to be met, to merely say that the outcome of treatment was measured in some way is not enough. The outcome measure itself must be evaluated to check its reliability. *Reliability* refers to the consistency of measurement. The reliability criterion here is satisfied only if the author of the study affirms that *evaluations were made of the outcome measure's reliability (for example, inter-rater agreement), and the author lists a numerical value of some kind for this measure of reliability*. Where multiple outcome criteria are used, reliability checks of any one of the major outcome criteria satisfy Criterion 13.
14. *Reliability measure has value greater than .70 or percent of rater agreement is greater than 70%:* The reliability coefficient in Criterion 13 is .70 or greater. Reliability coefficients typically range from -1 (perfect disagreement), through 0 (no pattern of agreement or disagreement), to 1 (perfect agreement).
15. *Those rating outcome rated it blind:* This criterion concerns the way bias can enter into measurement if the person measuring outcome knows whether the subject being measure is from a treatment or control group, or, worse, the person measuring outcome is in a position to determine the outcome measure. *Give the points for this criterion only if the person conducting the outcome measuring did not know which subjects were in treatment or control groups*.
16. *Treatment outcome was measured after treatment was completed:* *At least one outcome measure was obtained after treatment was completed*. Outcome measure both during treatment and after treatment is sufficient to meet this criterion.
17. *Test of statistical significance was made and $p < .05$:* Test of statistical significance are generally referred to by phrases such as "differences between treatment groups were significant at the .05 level" or "results show statistical significance for..." *Statistical significance* refers to the probability of obtaining an observed difference between treatment or control groups as great as or greater than by chance alone. Give credit for meeting this criterion only if the author refers to a test of statistical significance for a

EVIDENCE BASED PRACTICE
Evaluating the Quality of Evidence
Quality of Study Rating Form (QSRF)

major outcome variable naming the statistical procedure (e.g., analysis of variance, chi-square, *t* test) and gives a *p* value, for example $p < .05$, and the *p* value is equal to or smaller than .05.

18. *Follow-up was greater than 75%*: The proportion of subjects successfully followed up refers to the number contracted to measure outcome compared with the number who began the study. Ideally, the two should be the same (100% followed up). To compare the proportion followed up for each group studies (i.e., treatment group(s), control group), determine the number of subjects who initially entered the study in the group and determine the number successfully followed up. (If there is more than one follow-up period, use the longest one). Then, for each group, divide the number successfully followed up by the number who began in each group and multiply each quotient by 100. If the *smallest* of these percentages exceeds 75%, then the study meets this criterion.
19. *Total quality point (TQP) (add 1-18)*: Simply add the point values for Criteria 1-18 and record the value in Box 19. This value will range between 0 and 100.
20. *Effect size (ES1) or standardized mean difference* (magnitude of difference between groups in standard deviation units) calculated by the following:

$$ES1 = \frac{\text{mean of treatment} - \text{mean of control or alternate treatment}}{\text{standard deviation of control or alternate treatment}}$$

This formula is for computing ES1 (standardized mean difference or difference in standard deviation units) when outcome means of treatment and control groups are given. To compute effect size from information presented in a study's report, select two means to compare; for example, outcome might be a mean of a treatment group compared with a mean of a nontreatment control group. Subtract the mean of the second group from the mean of the first group and divide this difference by the standard deviation of the second group. ES1 may be a negative or positive number. If the outcome measure's score gets greater as client outcome improves and ES1 is positive, then the treatment has had a positive effect, proportionate to the size of ES1. In this case, if ES1 is negative, then the treatment harms.

21. *Effect size (ES2) or absolute risk reduction*:

$$ES2 = [(\text{number improved in treatment} \div \text{total number in treatment group}) \times 100] - [(\text{number improved in alternate treatment or control} \div \text{total number in alternate treatment or control}) \times 100]$$

Absolute risk reduction (ES2) refers to the event rate in treatment relative to the event rate in the control group. Assume that you are comparing the proportion in a treatment group who are improved against the proportion in the control group who are improved.

22. *Effect size (ES3) or number needed to treat*:

$$NNT = \frac{100}{ES2}$$

NNT (ES3) is the number of clients that a clinician must treat with the experimental treatment in order to create a good outcome or to prevent one bad outcome in comparison to the control treatment. If controls do better, then this number is the number needed to harm.

Impact of Parent Training on Family Functioning

Jerome F. Adams

ABSTRACT. This study examined outcomes of a STEP parent training program in several areas of family functioning identified by the so-called McMaster model of family health. The study found that participants who completed the parent training did report more improvements in general family functioning than those families whose children received routine mental health services. More specific improvements in family relationships were reported in the areas of problem solving, communication, affective responsiveness, and behavior control. Some differential effects were found for children older than 10 years of age. These improvements were also evaluated to determine their clinical significance, by examining how many families moved from problematic to healthy functioning. Implications of the findings are discussed. [Article copies available for a fee from The Haworth Document Delivery Service: 1-800-342-9678. E-mail address: <getinfo@haworthpressinc.com> Website: <<http://www.HaworthPress.com>> © 2001 by The Haworth Press, Inc. All rights reserved.]

KEYWORDS. Parent training, parenting skills, family functioning, psychotherapy outcomes, evaluation

The success of parent training for the reduction of child behavior problems has been well-documented (O'Dell, 1985; Kazdin, 1987, 1991; Webster-Stratton, 1985, 1991; Serketich, & Dumas, 1996). Although these programs have focused primarily on reducing antisocial behaviors such as noncompliance, temper tantrums, defiance, and

Jerome F. Adams, PhD, is Associate Professor of Family Therapy, Department of Human Development and Family Studies, The University of Rhode Island, Kingston, RI 02881.

Child & Family Behavior Therapy, Vol. 23(1) 2001
© 2001 by The Haworth Press, Inc. All rights reserved.

1.3

1.1

1.2

5

aggressiveness (Forehand, & McMahon, 1981; Patterson & Narrett, 1990; Emery, Binkoff, Houts, & Carr, 1983; Lutzker, McGimsey, McRae, & Campbell, 1983), these studies also report positive findings for parents. Benefits for parents include increased confidence in child management, a more accepting attitude toward their child, insight into the possible causes of their child's behavior, and improved self-esteem, (Williams, Omizo, & Abrams, 1984; Pehrson & Robinson, 1990; Mullin, Quigley, & Glanville, 1994).

Researchers have also investigated the efficacy of also treating additional family problems together with parent training. Adjunct therapies such as parent enhancement therapy (Griest, Forehand, Rogers, Breiner, Furey, & Williams, 1982), partner-support training (Dadds, Schwartz, & Sanders, 1987), and social learning family therapy (Sayger, Horne, & Glaser, 1993) have been shown to produce greater improvement in child behavior than parent training alone. These differences seem to be especially pronounced at follow-up assessments, suggesting that adjunct therapies may enhance the maintenance of gains made during parent training. Pfiffner, Jouriles, Brown, Etscheidt, and Kelly (1990), for example, report that single-parent families benefit more from parent training when general social problem-solving skills training focusing on current problems other than child management addressed directly during treatment.

Relatively little attention, however, has been given to systematically assessing the impact of child management training on other areas of family functioning. Studies that address this question provide limited evidence (Griest & Wells, 1983). Thompson, Ruma, Schuchmann, & Burke (1996) found a general improvement in satisfaction with family relationships but recommended a broader investigation of other family functioning variables known to promote positive outcomes.

The main purpose of this study was to determine the effects of parental participation in a highly structured Systematic Training for Effective Parenting (STEP) program including areas of family functioning other than child management. The study was also designed to evaluate two related research questions. The first concerns the clinical significance of the findings. For some time clinical investigators have questioned the validity of statistical significance without clinical significance of treatment effects (e.g., Kazdin, 1987, 1977). Clinical significance is broadly defined as improvement from the clinical to normal ranges during treatment. One method of evaluation concludes that

clinically significant gains are made when scores in the clinical range at pretest are in the normal range at posttest. More recently, however, two investigators have developed methods to rule out measurement error and test for clinical significance of therapeutic change (Jacobson, & Truax, 1991). Using this methodology, gains made by each subject are observed to determine if there is movement from the clinical to the normal range and whether this change is statistically reliable. Resulting data include rates of clinically significant change or recovery. With this more rigorous definition of clinical significance, can parent training produce such changes in overall family functioning?

The second question pertains to the ages of the target children identified by the parents. Previous research suggests differential effectiveness of parent training programs depending on the child's age (Dishion, & Patterson, 1992; Ruma, Burke, & Thompson, 1996). Families may be stressed in different ways depending on the age of the child and the nature of the behavior problems they present. This study sought to determine if family functioning were affected in different areas depending on the age of the child and whether these areas of functioning improved with increased skill in child management.

METHOD

Procedure

The STEP parenting groups were conducted at two outpatient mental health clinics serving a working class population. Participation was voluntary. Data from the two sites were pooled because of their comparability on demographic and dependent variables at pretest.

The research design is a pretest/posttest model with a parenting treatment group (PT) and no-parenting treatment (NTC) comparison group. Parents whose children were in treatment were recruited from each clinic to enroll in the STEP program. Parents were randomly assigned to one of the two treatment conditions. Once a sufficient number of parents had been recruited a parenting treatment group was established. Each parent was asked to specify the child living at home with whom he or she would apply the skills acquired during the training period (usually the child in treatment). Within two weeks of the start of each group parents completed the assessment measures.

parent training ⊕
m.h. services
vs.
m.h. services
alone

3

2.1

6

9

The comparison group consisted of parents not randomly assigned to the parenting groups. These parents did not participate in the study as PT parents but did have access to clinical services as requested. Comparison group parents completed the pretest measures in the same manner as the treatment group participants. The treatment group parents attended 4 hour weekly STEP group meetings for 8 weeks and were posttested during the final session with the same instruments that were administered prior to the initial session.

10

5

10

Subjects

Thirty-nine parents constituted the PT group. Eleven of 50 parents referred dropped out before the completion of the program. The NTC comparison group was composed of 35 parents. Fifteen of the original 50 parents assessed failed to complete the assessments in this group. The treatment groups were comparable in terms of gender and age of the identified target child, as well as parent demographic characteristics—parenting status, age, gender, and education level. Most of the parents were married females who had high school education (see Table 1.)

2.3

11

2.7

18

Parents were asked to identify a target child as a focus for the parenting skills training. This was defined as the child living at home with the most severe behavior problems. If more than one child in the home had equally severe behavior problems, parents were asked to select the oldest of these children (not over 18). Sixty-two percent of the target children were male. Target children ranged in age from 3-16 years with a median age of ten years.

MEASURES

McMaster Model of Family Functioning

Family functioning is conceptualized using the McMaster model (see Epstein, Bishop, and Levin, 1978), which emphasizes a functional approach to understanding how and whether families accomplish basic tasks of daily life. The McMaster model has proven useful for examining the course of individuals' behavior within the contexts of their family situations (Forman, & Hagan, 1984; Fristad, 1989; Grote-

TABLE 1. Family Demographics by Group

	PT N = 39	Control N = 35	F	η^2
Parent				
Age				
M	32.66	34.95	.761	
SD	6.35	9.42		
Family Composition				
1 Parent	44%	48%		.039
2 Parent	56%	52%		
Gender				
Female	77%	62%		1.15
Education				
M	12.11	12.37		
SD	1.45	3.22	.095	
Target Child				
Age				
M	9.78	9.95	.024	
SD	3.95	3.07		
Gender				
Male	62%	52%		.749

vant, & Carlson, 1989; Miller et al. 1985; Miller et al. 1994; Miller, Kabacoff, Keitner, Epstein, & Bishop, 1986).

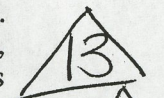
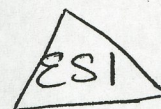
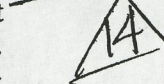
The model postulates six domains of family functioning: (a) Problem Solving is the family's ability to resolve both instrumental and affective issues to the level that maintains effective family functioning. (b) Communication is the relative clarity and directness of messages in both instrumental and affective exchanges of information among family members. (c) Roles encompasses the repetitive patterns of behavior by which family members function, including meeting basic needs, designation and responsibility for household tasks, as well as maintaining appropriate family boundaries and providing nurturance to family members. (d) Affective Responsiveness involves an individual family member's ability to express and experience appropriate range, quantity, and quality of feeling. (e) Affective Involvement focuses on the extent to which family members are interested and concerned, and value each other. (f) Behavior Control focuses on the rules and standards for behavior maintained by the family for all its members. (g) Overall family functioning is a global assessment of the family's ability to accomplish basic everyday tasks across domains.

← OFL's particular interest!

This includes information from each domain but not necessarily a linear combination of functioning in the specific areas.

Family Assessment Device (FAD)

The FAD (Epstein, Baldwin, & Bishop, 1983) is a 60-item self-report questionnaire based on the McMaster model yielding scores on each of the six dimensions of family functioning as well as a General Functioning score. The General Functioning scale is a composite scale designed to assess the overall health/pathology of the family. It consists of 12 items (one from Problem Solving, four from Communication, two from Roles, one from Affective Responsiveness, three from Affective Involvement and one from Behavior Control). Scores on the FAD range from 1 (very healthy) to 4 (very unhealthy), with questions worded to emphasize both positive and negative family functioning. Reliability and validity estimates of the FAD are adequate (Epstein, Baldwin, & Bishop, 1983; Miller et al. 1985). The FAD distinguishes between families rated by clinicians as healthy or unhealthy (Miller et al. 1985) on each dimension of the McMaster model (Keitner, Ryan, Miller, Epstein, & Bishop, 1989). The FAD is completed independently by the mothers and fathers on site under the supervision of the research assistant. The general family functioning score for each parent was used in the initial analyses, which with a clinical cutoff of 2 (i.e., scores of 2 or more indicate unhealthy functioning in the clinical range). Subsequent analyses examined family functioning on the six domains measured by the Family Assessment Device.

TREATMENT

Systematic Training for Effective Parenting or STEP (Dinkmeyer & McKay, 1976), is a highly structured parent training program. Parent groups are led by a trained professional who, with the assistance of supporting materials, reviews each of eight child management topics during a four hour weekly meeting. Topics are listed in chronological order in Table 2.



The program provides parents with three important sets of skills. The first set of skills offers the parents a new way of looking at their child's development. Instead of defining the child's behavior as de-

TABLE 2. Topics Discussed in STEP Meetings

-
- I. Understanding Children's Behavior and Misbehaviors
 - A. Four goals of misbehavior
 1. attention
 2. power
 3. revenge
 4. display of inadequacy
 - II. Understanding More About Your Child and Yourself as a Parent
 - A. Emotions
 - B. Family atmosphere and values
 - C. Sex roles
 - D. Methods of training
 - III. Encourage: Building Your Child's Confidence and Feelings of Worth
 - A. Competition between siblings
 - B. Overambition
 - C. Double standards
 - IV. Communication: How to Listen to Your Child
 - A. Effective listening
 - B. Reflective listening
 - V. Communication: Exploring Alternatives and Expressing Your Ideas and Feelings to Children
 - A. Concept of problem ownership
 - VI. Natural and Logical Consequences of Behavior
 - A. A method of discipline that develops responsibility
 - VII. Applying Natural and Logical Consequences to Other Concerns
 - A. Daily chores
 - B. Hygiene
 - VIII. Family Meetings and Leadership Skills
 - A. When to begin family meetings
 - B. Establishing meeting when only one parent is interested
 - C. Introducing family meetings to young children
-

fiant, this could be reinterpreted as developmentally appropriate. This enables the parent to develop more realistic expectations for their child. The second set of skills involves teaching parents new, more effective ways of communicating with and disciplining their children. A proposed additional benefit in helping a child to control his or her impulses through effective discipline is that parents become more aware of, and thus often more in control of, their own impulses. Finally, the program emphasizes the importance of decreasing a parent's sense of isolation. By sharing experiences in a group, parents become better aware of their own motives and needs, and learn that they are not alone. In addition, the sequential nature of the program provides a structured framework to help parents absorb newly presented skills and ideas as well as to build on to previously acquired ones.

Specific skills or content may be emphasized or reinforced through the repetition of a particular lesson or part of a lesson (Hitchcock, 1987).

RESULTS

Pretraining Comparisons

Pretest data for the treatment groups were analyzed using ANOVAs and chi-square analyses to test for equivalence of groups prior to treatment. There were no significant differences on demographic variables or pretest dependent measures. There were also no differences in the proportion of families included in the study from any of the respective clinics. Next, those who dropped out of the study were compared to those who completed the program. There were no significant differences on demographic variables or pretest dependent measures. This suggests that attrition is not a significant factor in the interpretation or generalization of results. Because of the small sample size, two groups (over 10, under 10) were coded for analysis based on the distribution and median age of the target child. Two (Group) \times 2 (Age) analyses of covariance (ANCOVAs) along with planned interactions were used to evaluate the effects of the STEP program on the primary outcome measures with the PRE score for each posttest measure used as the covariate. Reported means are adjusted for the PRE covariate.

Statistical Significance of Treatment Effects

Because General Functioning is a composite score of all areas of family functioning this was assessed first to determine treatment effects. Posttest analysis reveal a significant between-groups difference on this variable. Subsequent analyses were then done on each of the FAD variables. Means and standard deviations are reported in Table 3.

On the posttest analyses of covariance, parents in the STEP group reported significantly better scores than the comparison group in the areas of problem solving, communication, affective responsiveness, and behavior control. The FAD profiles for each treatment group are presented in Figure 1.

The adjusted means of the posttest scores were then examined to determine the effects for age of the target child. FAD profiles for child

(2.4)

(2.7)

Again, significant attrition

(2.3)

(3)

TABLE 3. Adjusted means and standard deviations for post-test scores on FAD family variables

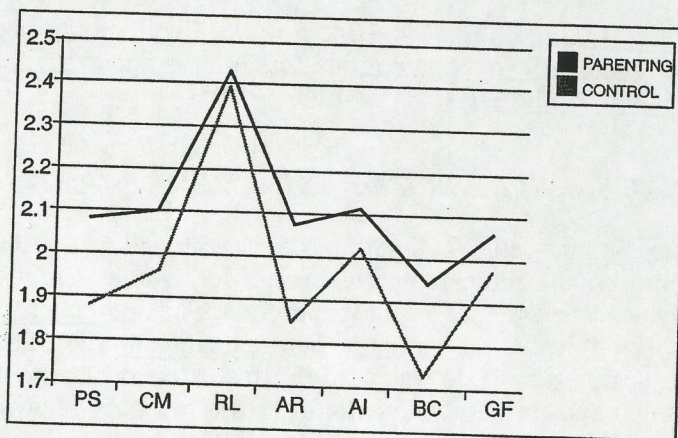
Variable	Parent Training (PT) N = 39		No-Treatment Comparison (NTC) N = 35		F
	Mean	S.D.	Mean	S.D.	
General Functioning	1.98	.59	2.06	.51	6.19*
Problem Solving	1.88	.51	2.08	.51	4.06*
Communication	1.96	.44	2.10	.43	8.12**
Roles	2.39	.34	2.43	.42	1.97
Affective Responsiveness	1.85	.68	2.08	.51	5.48**
Affective Involvement	2.02	.43	2.11	.44	.72
Behavior Control	1.73	.44	1.95	.48	4.03*

17

ES1
ES2
ES3

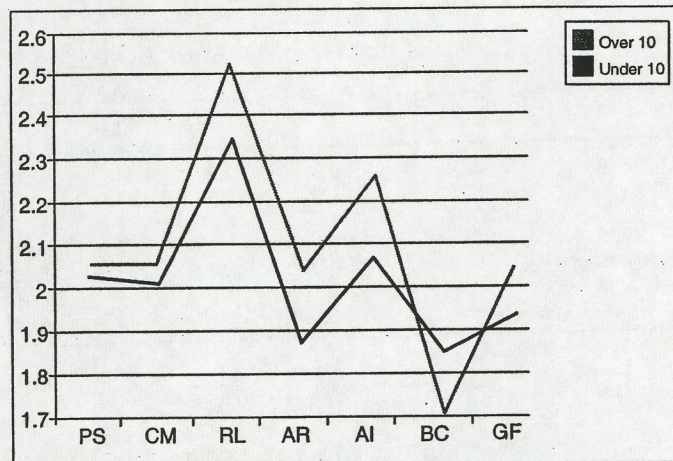
*p < .05
**p < .01

FIGURE 1. FAD scores for parent training and no treatment control groups.



age groups are presented in Figure 2. Analysis revealed no significant main effects for child age on any of the FAD variables even though the trend is for parents of children over 10 to report more family distress in all areas. The one exception is on the variable of behavior control, in which parents report more distress for children under 10. There are significant interaction effects on two of the FAD scores. Significantly healthier scores were found in the treatment group for children over 10.

FIGURE 2. FAD scores for target children over and under 10 years of age.



on the variable Roles, $F(1,64) = 4.34, p < .05$. Significantly better scores were found in the treatment group for children under 10 on the variable Behavior Control, $F(1,64) = 4.98, p < .05$.

CLINICAL SIGNIFICANCE OF TREATMENT EFFECTS

3
 In the current study clinical significance was defined as a reliable improvement from the clinical to normal range during treatment. Clinical significance was assessed using two criteria. First, parents had to move from the clinical to normal range during treatment. The General Functioning scale of the FAD was used for this assessment primarily because this composite family functioning scale has been shown to distinguish reliably normal and clinical populations. A cutoff Score of 2 and above defined the clinical range (Miller et al. 1985; Keitner, Ryan, Miller, Epstein, & Bishop, 1990). Second, pre-post gain scores were utilized to compute the change in functioning during the eight week treatment program. This change had to be equal to or greater than the reliable change index (Jacobson & Truax, 1991) in order to be considered a significant clinical recovery. Only the parents who scored in the clinical range before treatment were included in these analyses. In the NTC group, 23 (66%) of the 35 parents were in the clinical

clinical significance: moving from a clinical to a normal range

range at pretest (mean = 2.38; sd = .308). In the PT group, 21 (54%) of the 39 parents were in the clinical range at pretest (mean = 2.43; sd = .487). Twelve percent of these NTC parents and 38% of these PT parents made reliable improvements with recovery. This difference in recovery rates between the two groups is statistically significant, ($\chi^2(1, n = 44) = 5.40, p = .02$). These results should be interpreted with caution due to the fact that only 44 of the 74 subjects were included.

DISCUSSION

This study found that parents who completed the STEP parent training program reported healthier family functioning than those families whose children received routine mental health services in the areas of problem solving, communication, affective responsiveness, and behavior control. In addition, this healthier family functioning was found to be clinically as well as statistically significant.

The current study has some important limitations. First, the primary outcome measure relies exclusively on a self-report questionnaire from the parents. There were no direct measures of the parents' behavior with their children. It is possible that the improvements reflect what these parents say and not what they actually do in their families. Second, the improvements in family relationships may be unique to the relatively comprehensive nature of the STEP parenting program. Behavioral parent training programs which focus more tightly on the antecedents and consequences of child misbehavior (what Kazdin, 1997, refers to as the ABCs) may not generate such a generalized improvement in family relationships. Third, the STEP program was powerful enough to generate clinical recovery in only 38% of families in the treatment group. It is difficult to know how to interpret this finding, although other studies report recovery rates in this relatively low range (Thompson et al. 1996). And last, we were unable to determine the durability of treatment effects since there was no follow-up assessment.

Regarding the effects of the child's age on treatment outcome, this study appears to be consistent with other research suggesting poorer outcomes with older children (Kazdin, 1997; Dishion, & Patterson, 1992). Although not a statistically significant finding, parents in this study reported higher levels of dysfunction in all areas for children

over 10 years of age, with one exception. Parents reported more difficulty with Behavior Control when the target child is under 10 years of age. There also seems to be differential effects in specific areas depending on the age of the child. Parent training families report significantly better functioning in Roles when the child is over 10 years; better functioning in Behavior Control occurs when the child is under 10 years of age. Again, we need to be cautious in our interpretation of these findings. Age may be a proxy variable for problem severity. There is evidence (Ruma et al., 1996) that adolescents respond less well to parent training because they are more severely and chronically impaired on referral than pre-adolescents. Once severity is controlled, age does not influence outcome. Because our study did not assess problem severity we were unable to determine its impact on family relationships directly.

On the positive side, even though this exploratory study does not adhere to the demanding methodological standards set for clinical research (Peterson & Bell-Dolan, 1995) the results can still be used to advance practical knowledge about treating families. There are several strengths to the current study. First, the results indicate that the STEP parenting program generates benefits for families in areas of functioning other than those directly involved in child management when compared to a clinical control condition. Few studies have assessed the impact of parent training on broader areas of family functioning in this way. In addition, three aspects of this study are likely to make the results more meaningful to practitioners in the field. Reports of clinical recovery rates allow clinicians to determine if families reporting improvements greater than chance are functioning normally after treatment. The parenting treatment group is drawn from an outpatient clinical population, which makes it more like samples typically faced by clinicians. The comparison group includes families in which a child is currently receiving routine outpatient individually focused services.

REFERENCES

- Dadds, M. R., Schwartz, S., & Sanders, M. R. (1987). Marital discord and treatment outcome in behavioral treatment of child conduct disorders. *Journal of Consulting and Clinical Psychology, 55*, 396-403.
- Dinkmeyer, D., & McKay, G. D. (1976). *Systematic training for effective parenting*. Circle Pines, MN: American Guidance Services.
- Dishion, T. J., & Patterson, G. R. (1992). Age effects in parent training outcome. *Behavior Therapy, 23*, 719-729.

- Emery, R. E., Binkoff, J. A., Houts, A. C., & Carr, E. G. (1983). Children as independent variables: Some clinical implications of child-effects. *Behavior Therapy, 14*, 398-412.
- Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy, 9*, 171-180.
- Epstein, N. B., Bishop, D. S. & Levin, S. (1978). The McMaster model of family functioning. *Journal of Marriage and Family Counseling, 9*, 19-31.
- Forehand, R., Griest, D. L., Wells, K., & McMahan, R. J. (1982). Side effects of parent counseling on marital satisfaction. *Journal of Counseling Psychology, 29*, 104-107.
- Forehand, R., & McMahan, R. J. (1981). *Helping the noncompliant child: A clinician's guide to parent training*. New York: Guilford Press.
- Forman, B. D., & Hagan, B. J. (1984). Measures for evaluating total family functioning. *Family Therapy, 11*, 1-36.
- Fristad, M. A., (1989). A comparison of the McMaster and complex family assessment instruments. *Journal of Marital and Family Therapy, 15*, 259-269.
- Griest, D. L. & Wells, K. C. (1983). Behavioral family therapy with conduct disorders in children. *Behavior Therapy, 14*, 37-53.
- Griest, D. L., Forehand, R., Rogers, T., Breiner, J., Furey, W., & Williams, C. A. (1982). Effects of parent enhancement therapy on the treatment outcome and generalization of a parent training program. *Behaviour Research and Therapy, 20*, 429-436.
- Grotevant, H. D., & Carlson, C. I. (1989). *Family assessment: A guide to methods and measurements*. New York: Guilford Press
- Hitchcock, R. A. (1987). Understanding physical abuse as a lifestyle. *Individual Psychology, 43*, 51-55.
- 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.
- Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification, 1*, 427-451.
- Kazdin, A. E. (1987). Treatment of antisocial behavior in children: Current status and future directions. *Psychological Bulletin, 102*, 187-203.
- Kazdin, A. E. (1991). Effectiveness of psychotherapy with children and adolescents. *Journal of Consulting and Clinical Psychology, 59*, 785-798.
- Kazdin, A. E. (1997). Parent management training: Evidence, outcomes and issues. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(10), 1349(8).
- Keitner, G. I., Ryan, Miller, I. W., Epstein, N. B., & Bishop, D. S. (1989). Compounded depression and family functioning. *Comprehensive Psychiatry, 30*, 512-521.
- Lutzker, J. R., McGimsey, J. F., McRae, S., & Campbell, R. V. (1983). Behavioral parent training: There's so much more to do. *The Behavior Therapist, 6*, 110-112.
- Miller, I. W., Epstein, N. B., Bishop, D. S., Keitner, G. I., Baldwin, L. M., & van der Spuy, H. K. J. (1985). The McMaster Family Assessment Device: Reliability and validity. *Journal of Marital and Family Therapy, 11*, 345-356.

- Miller, I. W., Kabacoff, R. I., Epstein, N. B., Bishop, D. S., Keitner, G. I., Baldwin, L. M., & van der Spuy, H. I. J. (1994). The development of a clinical rating scale for the McMaster model of family functioning. *Family Process, 33*, 53-69.
- Miller, I. W., Kabacoff, R. I., Keitner, G. I., Epstein, N. B., & Bishop, D. S. (1986). Family functioning in the families of psychiatric patients. *Comprehensive Psychiatry, 27*, 302-312.
- Mullin, E., Quigley, K., & Glanville, B. (1994). A controlled evaluation of the impact of a parent training programme on child behaviour and mothers' general well being. *Counselling Psychology Quarterly, 7*, 167-180.
- O'Dell, S. (1985). Progress in parent training. In M. Hersen, R. M. Eisler, & P.M. Miller (Eds.), *Progress in Behavior Modification, 57-108*. New York: Academic.
- Patterson, G. R., & Narrett, C. M. (1990). The development of a reliable and valid treatment program for aggressive young children. *International Journal of Mental Health, 19*, 19-25.
- Pehrson, K. L., & Robinson, C. C. (1990). Parent education: Does it make a difference. *Child Study Journal, 20*, 221-236.
- Peterson, L., & Bell-Dolan, D. (1995). Treatment outcome research in child psychology: Realistic coping with the "Ten commandments of methodology." *Journal of Clinical Child Psychology, 24*, 149-162.
- Pfiffner, L. J., Jouriles, E. N., Brown, M. M., Etscheidt, M. A., & Kelly, J. A. (1990). Effects of problem-solving therapy on outcomes of parent training for single-parent families. *Child & Family Behavior Therapy, 12*, 1-11.
- Ruma, P., Burke, R., & Thompson, R. W. (1996). Group parenting: Is it effective for children of all ages? *Behavior Therapy, 27*, 159-169.
- Sayger, T. V., Horne, A. M., & Glaser, B. (1993). Marital satisfaction and social learning family therapy for child conduct problems: Generalization of treatment effects. *Journal of Marital and Family Therapy, 19*, 393-402.
- Serketich, W., & Dumas, J. (1996). The effectiveness of behavioral training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy, 27*, 171-186.
- Thompson, R. W., Ruma, P., Schuchmann, L., & Burke, R. (1996). A cost-effectiveness evaluation of parent training. *Journal of child and Family Studies, 5*, 415-429.
- Webster-Stratton, C. (1985). Predictors of treatment outcome in parent training for conduct disordered children. *Behavior Therapy, 16*, 223-243.
- Webster-Stratton, C. (1991). Annotation: Strategies for helping families with conduct disordered children. *Journal of Child Psychology and Psychiatry, 32*, 1047-1062.
- Williams, R., Omizo, M., & Abrams, B. (1984). Effects of STEP on parental attitudes and locus of control of their learning disabled children. *The School Counselor, 4*, 126-133.

RECEIVED: 08/20/99

REVISED: 01/10/00

ACCEPTED: 01/28/00