Psychological Bulletin
1987, Vol. 101, No. 3, 428-442

0033-2909/87/$00.75

Evaluating the Effectiveness of Family Therapies: An Integrative Review and Analysis

Mark D. Hazelrigg, Harris M. Cooper, and Charles M. Borduin
University of Missouri—Columbia

A number of previous reviews of studies evaluating the effectiveness of family therapy have varied in their conclusions. This review is the first attempt at an integrated statistical analysis of family therapy effectiveness research. Family therapy was found to have positive effects compared with both no-treatment and alternative treatment controls, as measured by family interactions and behavior ratings. Follow-up data demonstrated that family therapy continued to show positive effects over time, but the effects were diminished and more variable than at posttreatment assessment. Future research should concentrate on comparative outcome studies with specific populations and with both legitimate alternative treatments and placebo controls. In addition, researchers need to describe in greater detail the family therapy procedures that are used. Outcome measures should represent multiple vantage points and should include measures of family interactions, preferably using both self-report and observational methods.

Family therapy is the generic name for an extensive and heterogeneous group of treatment approaches that have emerged during the past 3 decades. The emergence of these treatment approaches reflects a major paradigmatic shift in the conceptualization of behavioral and mental disorders. This shift is from an analytic and reductionistic epistemology to one that is systemic and holistic (Hoffman, 1981; Keeney, 1983). Operating within this system paradigm, most family therapists share the assumption that problematic individual behaviors and symptoms are intimately related to patterns of interaction between family members. Although family therapy usually (but not always) involves face-to-face work with two or more family members, the epistemological basis of family therapy is not linked to who is included in treatment but rather to how family therapists think about problems and the human condition (Gurman, Kniskern, & Pinsof, 1986). A family systems perspective proposes that a particular behavior problem or symptom must always be understood in terms of the social context in which it occurs. The family is viewed not simply as a collection of individuals but as a rule-governed system and an organized group that transcends the sum of its separate elements.

Within this young and developing field, a number of different approaches to intervening in the family system have been described. One major approach is based on Bowen's multigenerational systems theory (Bowen, 1978). In this approach, the family is defined as including all of the extended family members, and therapy focuses on the individual's role within the extended family network. A second major approach is structural family therapy (Minuchin, 1974). The therapist works primarily with the nuclear family, and therapy is aimed toward altering the structure of the interactions between various family members. A third approach is strategic in nature and concentrates on resolving the presenting problem by directing family members and other relevant individuals to perform various therapeutic tasks (Fisch, Weakland, & Segal, 1982; Haley, 1976). A fourth major approach to family therapy involves the application of behavioral or social learning principles to resolving family problems (Patterson, 1971, 1982). In most cases, these interventions include training parents to modify children's problem behaviors in the home. In each of these approaches to family therapy, multiple interventions are often used with a given family. Interventions may also change to meet new goals that emerge during the course of therapy (Gurman et al., 1986).

Although systems theory has revolutionized the way that family therapists think about families, there is a major controversy within the field concerning which family therapy approaches are most consistent with a systemic epistemology (Dell, 1982; Hoffman, 1981; Keeney & Sprenkle, 1982). Pragmatic approaches to family therapy are principally concerned with behavioral outcomes. The emphasis is on rapidly alleviating symptomatic behavior through the disruption of maladaptive interaction sequences or inappropriate cross-generational power coalitions in the family system. On the other hand, aesthetic approaches to family therapy (e.g., Bowen, 1978; Whitaker, 1976), arising from phenomenological, psychodynamic, existential, and systems perspectives, are more concerned with the larger gestalt or ecosystem of which symptoms are a part. Therapists embodying the aesthetic perspective view the presenting symptom or complaint as a "motor for growth" and argue that the immediate alleviation of symptoms may be unethical because it can preclude the opportunity for a long-term, holistic healing of self (Keeney & Sprenkle, 1982). Aesthetic therapists, also referred to as the "new epistemologists" (Hoffman, 1981), believe that clinicians should not be in any hurry for change and tend to view the therapy process as a jour-

We would like to extend our gratitude to Nancy E. Cantor and two anonymous reviewers for their thoughtful and helpful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Harris M. Cooper, Center for Research in Social Behavior, 111 East Stewart Road, University of Missouri, Columbia, Missouri 65211.

ney that provides an opportunity to facilitate both their own and their clients' personal growth. Aesthetic therapists contend that pragmatic approaches to family therapy are reductionistic in their emphasis on solving the presenting problem and, as such, fall outside the "true" systemic paradigm (Keeney & Sprenkle, 1982).

Although pragmatic and aesthetic approaches to family therapy are being used by growing numbers of mental health professionals, claims for their effectiveness vary and are not always based on empirical data. In addition, many of the empirical studies that do exist are flawed, yielding results of questionable validity. The most common flaw is the failure to include a control group with which a group of clients receiving family therapy can be compared (see Gurman & Kniskern, 1978; Slipp & Kressel, 1979). These studies base their conclusions almost entirely on within-group differences before and after family therapy intervention. Therefore, any within-group differences could be due to many factors other than family therapy, such as other historical occurrences, biological changes in clients, and testing effects (see Cook & Campbell, 1979, for general criticisms of one-group pretest–posttest designs).

The purposes of this article are to review the literature on family therapy outcome studies and to quantify the conclusions of past research by statistically summarizing the data provided by these studies (Cooper, 1984; Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985; Rosenthal, 1984). The procedures used in this review are an attempt to provide a first approximation for systematically assessing the effectiveness of family therapies.

Previous Research Reviews

The intense interest in the effectiveness of family therapies has spawned at least 10 published reviews in the past 15 years. For example, Gurman et al. (1986) listed 22 works that include reviews of family therapy outcome research (including multiple reviews by the same authors and unpublished works). The 10 published reviews that were selected for examination in this article should be representative of all reviews in this area. The findings of these reviews are summarized in Table 1. There is no clear consensus among reviewers regarding the effectiveness of family therapy. Specifically, two reviewers concluded that family therapy is generally more effective than alternative treatments (Borduin, Henggeler, Hanson, & Harbin, 1982; Ulrici, 1983). Both of these reviews were limited to studies of adolescents, which is one possible reason for the more positive conclusions. Ulrici (1983) stated, "Family interventions have been generally successful . . . when compared to no-treatment and alternative treatment controls" (p. 33). However, other reviewers with positive conclusions made more tentative statements (DeWitt, 1978; Gurman & Kniskern, 1978; Massie & Beels, 1972; Masten, 1979; Paquin, 1977; Slipp & Kressel, 1979). For example, Gurman and Kniskern (1978) stated, "Family therapy is at least as effective and possibly more effective than individual therapy" (p. 883). Furthermore, two reviewers reached no overall conclusions because of the inconsistent findings and methodological problems in the studies they reviewed (Russell, Olson, Sprenkle, & Atilano, 1983; Wells & Dezen, 1978). Wells and Dezen (1978) stated, "Numerous methodological and practical difficulties beset the current body of family therapy research. Alone, or in combination, these have the effect of seriously weakening the conclusions that might otherwise have been drawn from the work" (p. 267).

There is consensus among the reviewers, however, that many studies of the outcome of family therapy have methodological flaws. The most commonly noted problems are (a) absence or inadequacy of control groups, (b) inconsistent measures of outcome, (c) lack of or insufficient delay before follow-up assessment, and (d) unclear and inconsistent theoretical basis for treatment. Although it is difficult to assess the results of poorly designed studies, when there are very few studies available every bit of data seems important. Some reviewers, therefore, have avoided rejecting studies of poorer quality to maintain a broad scope for their reviews.

Studies of family therapy effectiveness also vary in the number and type of subjects involved. Family therapy has been studied with families where children, adolescents, or one of the parents was identified as the patient. Studies also vary in the size of the sample, ranging from 3 to 4,303 patients. Finally, studies vary in terms of the training level of therapists, the type of treatment facility, and the length of therapy.

In sum, the search for an assessment of family therapy effectiveness has been frustrated by a research base that past reviewers have found difficult to synthesize. The problems of previous reviews stem in part from the traditional methods of research reviewing that were used. The traditional narrative method of summary and integration does not provide efficient mechanisms for combining the results of different studies or precisely weighting and assessing the effects of methodological variation. Nor do the results of traditional reviews allow easy comparison of results to other reviews concerning other types of therapy. Thus, the overall effectiveness of family therapy and its possible mediators has not been addressed precisely in past reviews.

Present Review

This review provides statistical tests of the overall conclusions reached about the effects of family therapy and expresses these conclusions in a form that is easily compared with results obtained in related fields. Other problems found in most previous reviews were also avoided. Standards for research design quality were established and studies failing to meet these criteria were excluded from the review. To ensure that a substantial data base was available for synthesis, we used exhaustive and systematic means to search the literature. This strategy resulted in the retrieval of more adequately performed published studies than have been included in past reviews.

However, the reader should not draw the conclusion, based on this selected sample, that the research on family therapy is in fact homogeneous in conceptualization and design. Rather, we have chosen studies for inclusion in the meta-analysis that permit the most justifiable inferences about family therapy and its relative effectiveness. We have left our analysis of the broader array of work in the area and of the practical realities of family therapy for narrative consideration.

Furthermore, in this review the effect of variations in design among studies was considered an empirical question. That is, differences in study methods were tested statistically to determine if they were associated with the effectiveness of family

therapy. Finally, studies were precisely weighted so that each study's contribution to the overall conclusions was reflective of its relative sample size.

Family Therapy as a Broad Conceptual Category

The theoretical stance of family therapists provides justification for combining disparate types of outcome studies under a unitary construct. By globally defining symptoms and adjustment difficulties as maladaptive patterns of interaction between people, family therapists often do not emphasize differences between types and severities of disorders. Family therapy is generally designed to alter the behavioral and interaction patterns of members in a system and does not directly focus treatment efforts on the individual identified as having the problem. Any study of family therapy outcome, therefore, is a study of how effectively the broad category *family therapy* changed the patterns of behavior within the family, which should then result in the resolution of the original problem. Although many different therapy techniques may be used, the goals of most family systems therapists are similar. According to the family systems theory, then, all family therapy outcome research is essentially studying the same thing—effectiveness with changing family interactions—and can reasonably be combined and synthesized.

There are, however, some important issues that should be borne in mind when considering our selection criteria and findings. First, there is a good deal of complexity that is hidden within definitions of family therapy. Although family therapies are generally designed to alter pertinent behavioral sequences or interaction patterns within family systems, there are many differences between family therapists in their approaches to systemic change. Indeed, the broad field of family therapy has been frequently classified in terms of different rival schools or camps, with each emphasizing different concepts and methods (see, e.g., Goldenberg & Goldenberg, 1985; Nichols, 1984). Within these schools, interventions may primarily involve work with individuals or couples (e.g., Bowen, 1978), the parental dyad (e.g., Patterson, 1982), the total household (e.g., Haley, 1976), or multiple ecological contexts including the family, work and academic settings, peer groups, and social service systems (e.g., Attneave, 1976; Henggeler, 1982; O'Connor & Lubin, 1984).

In spite of the differences in techniques used by various family therapy camps, there is common ground that justifies combining the results of studies from different perspectives. Strategic family therapists (e.g., Haley, 1976) and at least some behavioral family therapists (e.g., Alexander & Parsons, 1973) do recognize that the larger system, rather than the individual with the presenting problem, should be the unit that must change. However, in contrast to aesthetic therapists, others believe changes will be most effective through limiting the therapist's focus to the presenting problems (Keeney & Sprenkle, 1982). Pragmatic therapists argue there cannot be a nonsystemic intervention, because effecting stable changes in any part of a system requires adjustment throughout the system (Montalvo & Haley, 1973). The methods of intervening may differ among family therapists, but the overall goal is generally one of systemic change. Nevertheless, we recognize that in some of the studies in this review the therapists may have effected interactional change without a broad systemic vision.

One problem encountered in searching for and selecting studies for this review was that outcome studies of family therapy from an aesthetic perspective were not found. Thus, this review is composed of studies using a pragmatic approach to family therapy (primarily structural, strategic, and behavioral approaches), and long-term and intergenerational approaches are not represented. The lack of studies from an aesthetic perspective is due in part to publication lag, resulting in the omission of recently completed studies. However, a major factor is that aesthetically oriented therapists tend to be least likely to use traditional research methods, resulting in an overall dearth of research from this perspective (Gurman et al., 1986).

Another potential problem with the meta-analytic approach is that lumping several studies together may yield an overly general picture of how well family therapy works that is not applicable to specific techniques or situations. For example, different client populations may have different patterns in their family interactions (Hetherington, Stouwie, & Ridberg, 1971), and thus may respond differently to various family therapy approaches. At present, this point is untestable because although some theorists have developed ways to classify family interactions (e.g., Minuchin, 1974; Olson, Russell, & Sprenkle, 1979), family therapy outcome research does not report how the families treated fit into any classification schemes. Although the argument that different types of families will respond differently to various family therapy techniques may be valid, the extant research does not address the question.

Distinctions Among Studies

In this review, the results of the individual studies were grouped into various categories before any statistical procedures were used to combine the results. One distinction was the type of control group used in a study. Studies using control groups that consisted of alternative treatments and studies using control groups that recieved no treatment were analyzed separately. This was done because some treatment effects can be expected to occur in the alternative treatment groups, whereas no treatment effect should be found in the no-treatment conditions. Thus, the effects of family therapy should be greatest when compared with groups that have no treatment, whereas comparisons with alternative treatments examine the relative effectiveness of two therapies. When one study included both types of control groups, it was included in both analyses.

The studies were also distinguished according to the type of dependent variable used. Three categories of measurements were used often enough to allow separate analysis. These categories were measures of family interactions, behavior ratings of the identified patient, and recidivism. It should be noted that not all family therapists view symptom removal and prevention as constituting success in therapy. Indeed, as noted earlier, many family therapists view symptom removal as a beginning point for systemic change and not as an end point. Moreover, some families seek treatment for assistance in dealing with nonsymptomatic problems, such as reactions to developmental hurdles in the family life cycle (e.g., adolescent emancipation), stresses due to extrafamilial forces (e.g., unemployment), or idiosyncratic problems (e.g., physical illness).

Family interaction measures might be obtained by self-re-

Table 1
Summary of Previous Reviews of the Effectiveness of Family Therapy

Review	Number of studies	Conclusions	Noted methodological flaws
Borduin, Henggeler, Hanson, & Harbin (1982)	20	Family therapy generally more effective than alternate treatment.	Studies used different outcome measures. Follow-up intervals too short. Treatment procedures not well described.
DeWitt (1978)	31 (23 uncontrolled)	Family therapy better than no treatment, but relative effectiveness unclear compared to alternative treatments.	Therapists' characteristics and techniques not well described. Studies used different outcome measures.
Gurman & Kniskern (1978)	61 (38 uncontrolled)	Family therapy equal to or superior to alternative treatments. Family therapy at least as effective as individual therapy.	Studies used different outcome measures. Studies lack or use inadequate/ inappropriate control groups.
Massie & Beels (1972)	8 (7 uncontrolled)	Family therapy effective, especially for crisis resolution.	Studies used different outcome measures. Follow-up intervals too short. Techniques not well described. Theory vaguely described. Treated populations too heterogeneous. Family characteristics not classified.
Masten (1979)	14 (8 uncontrolled)	Family therapy effective in treating adolescents. No other conclusions drawn due to methodological flaws.	Lack of control groups Studies used single dependent variable.
Paquin (1977)	2	Trend favoring brief, directive therapy.	Follow-up intervals too short. Family characteristics not classified.
Russell, Olson, Sprenkle, & Atilano (1983)	9	No conclusions due to methodological flaws.	Few studies link family dynamics to treatment goals. Theory vaguely described.
Slipp & Kressel (1979)	4	Family therapy at least as effective as individual therapy.	Studies used different outcome measures. Lack of control groups. Studies used single dependent variable.
Ulrici (1983)	10	Family therapy more effective than alternative treatments. Behavior therapy most effective when whole family included.	Many variables not controlled such as therapist characteristics and treatment settings.
Wells & Dezen (1978)	19 (7 uncontrolled)	Findings inconsistent; no overall conclusions drawn.	Outcome measures vague and subjective. Family therapy confounded with other interventions. Nonrandom assignment to therapy groups. Techniques not well described. Studies used different outcome measures.

ports of the families' organization and relationships or by observation of the family by trained assistants. Observers could code the families' communication and behavior in a structured activity by rating the amount of conflict present or by recording the amount of participation of each member. Behavior ratings include measures of the identified patient's behavior, such as reports of the frequency of certain behaviors or evaluations of the patient's behavior. These ratings might be provided by various sources, such as parents, teachers, or research team observers. In this vein, several studies had a child's parent list problem

behaviors and monitor the frequency of those behaviors over a specific period of time. Finally, the recidivism rates used were measures of whether or not a patient returned to the referring agency with continuing problems.

A final distinction among studies concerned the timing of the assessment of treatment effects. Most studies assessed subjects immediately following treatment (posttreatment). However, some studies also assessed subjects again at a later date (follow-up) to determine if the effects of treatment were lasting changes. All results from posttreatment assessments were analyzed separately from any follow-up assessment results. Recidivism rates are also a form of follow-up and were also analyzed separately.

Literature Review Procedures

Literature Search

The studies reviewed were located from studies listed in *Psychological Abstracts* and *Educational Resources Information Center.* Both lists were obtained through a computer search using the keywords *family therapy* crossed with *outcomes, results, effects, evaluation,* or *impact.* The search included January 1966 to May 1984 and uncovered 281 articles. A reading of abstracts and full reports led to 15 studies that met the selection criteria for inclusion in the review.

In addition, the reference lists of previous reviews of family therapy outcome studies (see Table 1) were used to locate articles. Nine additional studies were found in this manner.

All of the studies were published in English language journals. No unpublished studies were included. This exclusion may be a source of bias and could influence the results, although the existence and direction of this influence cannot be predicted (Glass et al., 1981). The reviews of Gurman and Kniskern (1978) and Wells and Dezen (1978) did include some unpublished studies. However, these reviews did not address the issue of whether the unpublished studies were different from the published studies in either quality or direction of results.

Selection Criteria

For a study to be used in the review, it must have included families as the subject population, that is, a minimum of one parent and one child. Furthermore, at least one therapy process must have qualified as family therapy. Family therapy was defined as therapy using interventions intended to produce change in the families' interactions. Studies of marital therapy with no children involved were excluded.

Studies must also have included a control group in the experimental design. The control group could have received any of several alternative treatments or no treatment. Alternative treatments were defined as any intervention that was presented to the family as being therapeutic. Typically, the interventions in the control groups were focused on treating only the identified patient. Control groups that used placebo or attention interventions (e.g., bibliotherapy) were included as alternative treatments. When a study used two or more alternative treatment control groups, they were combined within the study to form a single alternative treatment group for these analyses.

Random assignment of patients to groups was not required. However, some attempt to equate the groups through subject matching or a statistical demonstration of the equivalence of the groups before treatment was required. Studies needed to have a minimum of five families in each treatment group.

A final requirement for inclusion was that studies thoroughly report results and the statistical analyses performed. This information was vital to compare the effects of family therapy across studies. Any study reporting either detailed analyses or enough raw data so that the needed statistics could be computed was included. Numerous studies reported adequate information for some measures but omitted detailed reports on other measures. In this review, only those dependent measures on which statistical tests were performed are included.

Search Outcome

A total of 20 studies contained in 24 journal articles were located that met these criteria. The reports of some studies can be found in several different journal articles that report different measures or lengths of follow-up times for the same groups of families. These reports were combined to obtain one measure of effect size for each group of subjects. At least 3 times the number of studies included in this review were not used because of flawed methodology, usually a lack of an adequate control group.

In light of the outlined criteria, several studies, both included and excluded, deserve special mention. One study, by Schubert and Miller (1981), was included even though it did not use random or controlled assignment to groups. It was included because the very large sample of clients was drawn from a relatively homogeneous population and the decisions about assignment to treatment conditions seemed to be based on unbiased criteria. Furthermore, the effect of family therapy was not the primary concern of the authors, who were focusing on the effect of social class within therapy types.

Several studies that included control groups were excluded for various reasons. Four studies did not provide reports of their data that were detailed enough to include in the statistical analyses (Douds, Engelsgjerd, & Collingwood, 1977; Hendricks, 1971; Pevsner, 1982; Wiltz & Patterson, 1974). Two studies used nonrandom assignment of subjects to groups without matching subjects on relevant variables (Bogart & French, 1978; Gould & Glick, 1977). Four studies used control groups consisting of clients who had dropped out of therapy prematurely but were still available for follow-up assessment (Gruher, 1979; Ostensen, 1981; Sigal, Barrs, & Doubilet, 1976; Weathers & Liberman, 1975). Clearly, families that successfully complete therapy cannot be compared with families that would be considered by many therapists as treatment failures.

One study (Szapocznik, Kurtines, Foote, Perez-Vidal, & Hervis, 1983) was excluded because the control group's treatment was intended to change family interactions and was labeled as family therapy even though the treatment involved only one person in the family.

Statistical Procedures

To synthesize the data, four statistics were calculated. First, the probability that the combined result from all studies was

Table 2
Studies Comparing Family Therapy With No-Treatment Controls Using Measures of Family Interactions Immediately After Treatment

Study	Patient population	Dependent measures	*N*	*d*	*Z*
Foster, Prinz, & O'Leary (1983)	Adolescents with behavior problems.	Self-report questionnaires. Experimenter ratings of recorded interactions.	18	.31	0.17
Garrigan & Bambrick (1975)	Children and adolescents with behavior problems.	Self-report questionnaires.	18	.24	0.40
Garrigan & Bambrick (1977, 1979, Study 1)	Children and adolescents with behavior problems.	Self-report questionnaires.	44	.44	0.94
Garrigan & Bambrick (1979, Study 2)	Children and adolescents with behavior problems.	Self-report questionnaires.	24	−.01	0.0
Hardcastle (1977)	Children with behavior problems.	Self-report questionnaires, experimenter ratings of recorded interactions.	28	.89	2.02
Katz, de Krasinski, Philip, & Wieser (1975)	Children, adolescents, and young adults with behavior problems.	Experimenter rating of recorded interactions.	11	.51	0.61
Reiter & Kilmann (1975)	Children with behavior problems.	Self-report questionnaires. Experimenter rating of recorded interactions.	24	.64	0.85

due to chance was calculated using the method of adding *Z* scores (Rosenthal, 1978). This was done by calculating a probability level for each study as a whole. When more than one outcome measure was reported in a study, the average of the *Z* scores associated with the *p* levels of relevant dependent variables was used. This ensured that each study contributed equally to the combined *p* level. These *Z* scores were summed across studies and divided by the square root of the number of studies. The resulting statistic is itself a *Z* score associated with an overall probability level indicating the likelihood that the combined results of the studies could have occurred by chance.

The average combined probability level was used to calculate the second statistic, termed the fail-safe *N* (Cooper, 1979). This is the number of studies that sum to a null effect that would have to be added to the known studies to raise the combined probability to nonsignificance. Rosenthal (1979) referred to this as the "tolerance for future null results" (p. 638). This statistic provides some measure of the stability of the findings in relation to the possible number of unretrieved studies.

Third, the effect size associated with each dependent variable was calculated. This review used the *d* index to measure effect sizes (Cohen, 1977). The *d* index is defined as the difference between the mean scores of two groups divided by the average or common standard deviation of the groups. This calculation results in a measure of the degree to which the two groups differ in terms of standard deviation units. For example, a *d* index of .25 indicates that the two group means were separated by one quarter of a standard deviation. In this review, the control group mean was always subtracted from the family therapy mean so that positive *d* indexes indicate that the family therapy group showed higher scores on measures of positive family interaction or patient behavior. The mean *d* index across studies was calculated by averaging the separate study *d* indexes after they had been weighted by the study sample size (Hedges & Olkin, 1985).

If multiple effects were reported within a study, they were averaged to yield a single overall *d* index. In cases where means and standard deviations were not reported, *d* index values were estimated from *F* values or *t* values (see Friedman, 1968). If an effect was reported as not significant and the *F* or *t* value was not reported, we assumed that an exact null effect was uncovered (i.e., $d = 0$). This convention probably causes a conservative bias in the results. That is, a lower magnitude of the *d*-index average value will be estimated than would be found if all studies included the necessary statistics.

Associated with a *d* index is a measure of distribution overlap, called U_3. U_3 tells the percentage of people in the lower meaned group who are surpassed by the average person in the higher meaned group. Thus, if $d = .25$, $U_3 = .60$, meaning the average patient in the family therapy condition surpassed 60% of the control condition patients.

The fourth statistic was a measure of the homogeneity of effect sizes in a group of studies (Rosenthal, 1984). This test determines whether the variability in a set of effect sizes is greater than would be expected by sampling variability alone. It results in a statistic that is distributed according to the standard chi-square distribution. If the chi-square is not significant, the reviewer can assume the effect indexes reported for the group of studies were measuring the same relation (i.e., were drawn from the same population). If the chi-square is signifi-

Table 3
Studies Comparing Family Therapy With No-Treatment Controls Using Behavior Ratings Immediately After Treatment

Study	Patient population	Dependent measure	*N*	*d*	*Z*
Fischer, Anderson, Arveson, & Brown (1978)	Children with behavior problems.	Parent evaluations of child behavior.	37	1.02	3.29
Garrigan & Bambrick (1975)	Children and adolescents with behavior problems.	Behavior checklist of teachers.	18	.33	1.15
Garrigan & Bambrick (1977, 1979, Study 1)	Children and adolescents with behavior problems.	Parent and teacher ratings of child behavior.	44	.40	0.94
Garrigan & Bambrick (1979, Study 2)	Children and adolescents with behavior problems.	Maternal ratings of child behavior.	24	−.31	−0.70
Martin (1977)	Children with behavior problems.	Maternal monitoring of child behavior.	43	.77	2.43
Stover & Guerney (1967)	Children with behavior problems.	Experimenter ratings of child behavior from recorded activity.	28	.38	1.45

cant, the reviewer should try to determine which studies (or effect sizes) might be included in further subsets of the studies.

In six of the reviewed studies, the raw data on recidivism was either presented or could be calculated from percentages of recidivating patients and the sample size. Using this raw data we calculated a single 2 × 2 chi-square involving the factors of (a) the number of clients rehospitalized or not rehospitalized and (b) whether they were treated with family therapy or an alternative therapy. An additional 2 × 6 chi-square was calculated involving the factors of (a) recidivism rate per 100 observations and (b) the study reporting the recidivism rate. This analysis tested whether recidivism rates were generalizable over all of the studies.

Results

Family Therapy Compared With No Treatment

Ten studies meeting the selection criteria included measures of the effect of family therapy compared with a no-treatment control group. Seven of the studies reported measures of family interactions and six studies reported ratings of the behavior of the identified patient.

Family interaction measures. The seven studies including family interaction measures are listed in Table 2. The combined *Z* score of these results yielded a value of 1.88, with an associated $p < .03$, one-tailed.

The fail-safe *N* for the results of family interaction measures comparing family therapy with no treatment was 3. This means three studies with a null-summing effect would have to be added to the seven reviewed studies before the cumulative result became nonsignificant. Thus, the findings cannot be viewed as stable and resistant to the effects of new or unretrieved studies that find null results.

The effect sizes for the studies in Table 2 ranged from $d = -.01$ to $d = .89$. Only one study reported results favoring the no-treatment group (Garrigan & Bambrick, 1979), and this result was a near-zero effect. All other studies reported results in a positive direction. The weighted mean effect size for all seven studies was $d = .45$. This effect size is associated with a U_3 value of 67.3%, which means that 67.3% of the people in the no-treatment group showed less favorable family interaction patterns than did the average patient in the family therapy group.

Analysis of the homogeneity of effect sizes revealed a nonsignificant chi-square, $\chi^2(6, N = 7) = 2.94$. Thus, the effect sizes were shown to be homogeneous.

Behavior ratings. The six studies that obtained behavior ratings of patients are listed in Table 3. The combined *Z* score of these results yielded a value of 3.49 with an associated $p < .0002$.

The fail-safe *N* for the behavior ratings comparing family therapy to no treatment was 26. This means 26 studies with a null-summing effect would have to be added to the 6 reviewed studies before the cumulative result became nonsignificant. Thus, the findings can be viewed as stable and fairly resistant to the effects of new or unretrieved studies that find null results.

The effect sizes for the studies in Table 3 ranged from $d = -.31$ to $d = 1.02$. Only one study reported results favoring the no-treatment group (Garrigan & Bambrick, 1979). All other studies reported results in a positive direction. The weighted mean effect size for all six studies was $d = .50$. This effect size is associated with a U_3 value of 69.1%, which means that 69.1% of the people in the no-treatment group had poorer behavior ratings than did the average patient in the family therapy group. The reason behavior ratings had a more significant combined probability than family interaction ratings did, even though the effect sizes were similar, appears to be the larger sample sizes in the behavior measure studies.

Analysis of the homogeneity of effect sizes yielded a nonsignificant chi-square, $\chi^2(5, N = 6) = 6.57$. Thus, the effect sizes were shown to be homogeneous.

Table 4
Studies Comparing Family Therapy With Alternative Treatment Controls Using Measures of Family Interactions Immediately Following Treatment

Study	Patient population	Alternative treatment	Dependent measure	*N*	*d*	*Z*
Foster, Prinz, & O'Leary (1983)	Adolescents with behavior problems.	Education in problem-solving skills.	Self-report questionnaires, experimenter ratings of recorded interactions.	17	.31	0.50
Johnson & Maloney (1977)	Children (not psychotic or delinquent).	Individual therapy with various family members.	Experimenter ratings of recorded interactions.	28	1.16	2.61
Parsons & Alexander (1973)	Adolescents (juvenile offenders).	Individual therapy (client-centered and eclectic/dynamic).	Experimenter ratings of recorded interactions.	40	.44	0.93

Family Therapy Compared with Alternative Treatments

Eleven studies meeting the selection criteria included measures of the effects of family therapy compared with the effects of alternative treatments (e.g., individual or group therapy, or medication). Eight of these studies provided data comparing the groups immediately following treatment, and seven studies provided follow-up data. The follow-up data were analyzed separately for this review and will be discussed later.

Family interaction measures. Three studies used family interaction measures to compare the effects of family therapy and alternative treatments (see Table 4). The combined Z score of these results yielded a value of 2.33 with an associated $p < .01$.

The fail-safe N for these studies is 3. This category of studies cannot be considered resistant to the effect of new or unretrieved studies reporting null effects.

The effect sizes for these studies ranged from $d = .31$ to $d = 1.16$. All of the studies reported results in the direction favoring family therapy. The weighted mean effect size was $d = .65$. This effect size is associated with a U_3 value of 74.2%, which means that 74.2% of the people in the alternative treatment group had less favorable family interactions than did the average person in the family therapy group.

Analysis of the effect-size homogeneity revealed a nonsignificant chi-square, $\chi^2(2, N = 3) = 2.48$. Thus, these effect sizes were found to be homogeneous.

Behavior ratings. Four studies reported using behavior ratings of the identified patient to compare the effects of family therapy and alternative treatments (see Table 5). The combined Z score of these results yielded a value of 2.63 with an associated $p < .004$.

The fail-safe N for these studies is 7. This category of studies cannot be considered resistant to the effect of new or unretrieved studies reporting null effects.

The effect sizes for these studies ranged from $d = 0.0$ to $d = .90$. All of the studies reported results in the direction favoring family therapy. The weighted mean effect size was $d = .23$. This effect size is associated with a U_3 value of 60.0%, which means that 60.0% of the people in the alternative treatment group had poorer behavior ratings than did the average person in the family therapy group.

Analysis of homogeneity of the effect sizes yielded $\chi^2(3, N = 4) = 8.10$, $p < .05$. Thus, these effect sizes were found to be heterogeneous.

Follow-Up Data From Studies Comparing Family Therapy With No-Treatment Controls

Two studies comparing family therapy with no treatment obtained follow-up data. Garrigan and Bambrick (1977, 1979) used a follow-up period of 1 to 2 years. The outcome measures were employment or school status and court involvement. This study reported an average effect size of $d = 1.45$, with a Z score of 3.97, $p < .00005$. Klein, Alexander, and Parsons (1977) used a follow-up period of 2.5 to 3.5 years. The outcome measure was the rate at which siblings of the identified patient were referred to the court system. These authors reported results with an effect size of $d = .37$ and a Z score of .85, $p < .20$. Because of the small number of studies reporting follow-up data from no-treatment control groups, no analyses were performed to combine results.

Follow-Up Data From Studies Comparing Family Therapy With Alternative Treatment Controls

Seven studies obtained follow-up data comparing the effects of family therapy to alternative treatments. Six studies used recidivism rates at the end of the follow-up interval and four studies used other measures. The results of these four studies are summarized in Table 6.

The amount of time for follow-up assessments ranged from 6 to 8 weeks to over 3 years. The standard normal deviate of the measures was $Z = 0.97$, *ns*. The effects sizes for these studies ranged from -1.22 to .90 with a weighted mean of $d = .06$. This effect size is associated with a U_3 value of 52.0%, which means that 52.0% of the people in the alternative treatment group were surpassed by the average person in the family therapy group. The homogeneity analysis for these effects resulted in $\chi^2(3, N = 4) = 23.51$, $p < .005$, demonstrating that the results are heterogeneous.

Six studies obtained follow-up data that specified the number of patients in each therapy group returning to treatment during

Table 5
Studies Comparing Family Therapy With Alternative Treatment Controls Using Behavior Ratings Immediately Following Treatment

Study	Patient population	Alternative treatment	Dependent measures	*N*	*d*	*Z*
Christensen, Johnson, Phillips, & Glasgow (1980)	Children with behavior problems.	Bibliotherapy.	Experimenter and parent monitoring of child behavior.	28	.90	1.96
Langsley, Machotka, & Flomenhaft (1971) Langsley, Pittman, Machotka, & Flomenhaft (1968)	Adults referred for inpatient treatment.	Hospitalization.	Experimenter ratings of patients' level of functioning.	250	0.0	0.0
Schubert & Miller (1981)	Adults in inpatient treatment.	Individual therapy, group therapy, behavior modification, and medication.	Experimenter ratings of patients' level of functioning.	4303	.24	1.73
Stuart, Jayaratne, & Tripodi (1976)	Adolescents (juvenile offenders).	Group therapy.	Parent and teacher ratings of child behavior.	60	.52	1.57

the follow-up period (Alexander & Parsons, 1973, 6- to 18-month follow-up; Goldstein, Rodnick, Evans, May, & Steinberg, 1978, 6-month follow-up; Johnson, 1977, 2-year follow-up; Langsley, Machotka, & Flomenhaft, 1971, and Langsley, Pittman, Machotka, & Flomenhaft, 1968, 6-, 12-, and 18-month follow-up; Rittenhouse, 1970, 3-, 6-, and 12-month follow-up).

A total of 952 patients were included in the recidivism analysis. These data were used in a single chi-square analysis (see Table 7). Three studies (Langsley et al., 1971, 1968; Rittenhouse, 1970; Wellisch & Ro-Trock, 1980) reported recidivism rates separately for more than one follow-up period. Although it is possible that the same patients recidivated repeatedly and are included in each follow-up period, these analyses are based on a total across periods.

The overall $\chi^2(1, N = 952) = 46.18, p < .005$. In percentage terms, during the follow-up period 36% of the patients in family therapy returned to treatment whereas 58% of the alternative therapy patients returned to treatment. The effect size for the recidivism data was $d = .47$.

The second chi-square analysis determined whether the effect sizes were consistent across studies. This analysis yielded a significant chi-square value, $\chi^2(5, N = 512) = 25.73, p < .005$, showing that the effects were heterogeneous.

The recidivism data was also subdivided according to patient population. First, hospitalized patients only were included (omitting data from Alexander & Parsons, 1973, and Johnson, 1977, who measured the number of court referrals during the follow-up period). The chi-square value for hospitalized patients only was $\chi^2(1, N = 436) = 20.93, p < .0005$, with an effect size of $d = .41$. In percentage terms, during the follow-up period 28% of the patients in family therapy returned to the hospital whereas 48% of control patients required further hospitalization. The second analysis to determine whether the effects were consistent across studies also yielded a significant chi-square value, $\chi^2(3, N = 308) = 23.89, p < .005$, showing that these effects were heterogeneous.

Next, the three studies of adolescent subjects (and their families) were combined (Alexander & Parsons, 1973; Johnson, 1977; Wellisch & Ro-Trock, 1980). For recidivism among adolescents, $\chi^2(1, N = 494) = 21.76, p < .005$, with an effect size of $d = .43$. In percentage terms, during the follow-up period 45% of the adolescent patients in family therapy returned to treatment whereas 66% of the adolescents in alternative treatment required further treatment. The second chi-square analysis to determine whether the effect sizes were consistent across studies yielded a significant chi-square value, $\chi^2(2, N = 318) = 6.97, p < .05$, indicating that these effects were heterogeneous.

For the six studies reporting recidivism data, a Spearman rho correlation coefficient was calculated between the size of the effects and the length of the follow-up periods. The result was a negative correlation ($r = -.30$), indicating a trend for effect sizes to be smaller when measured at longer periods of follow-up. Thus, the difference in recidivism rates among studies is likely due to both the use of different alternative treatments and different follow-up intervals.

Random Versus Nonrandom Assignment

Finally, a comparison was made between studies using random assignment and studies using nonrandom assignment (usually matching designs). In studies comparing family therapy with no treatment and using family interaction measures, five studies used random assignment and had a weighted average effect size of $d = .44$. Two studies used nonrandom assignment, and these studies had a weighted average of $d = .45$. In studies using behavior ratings to compare family therapy with no treatment, four used random assignment, with $d = .38$, and two used nonrandom assignment, with $d = .79$. In studies comparing family therapy with alternative treatments, all studies using measures of family interactions used random assignment. Among studies using behavior ratings, three used random assignment, with $d = .18$, and two used nonrandom assignment, with $d = .24$.

Table 6
Studies Comparing Family Therapy With Alternative Treatment Controls Using Nonrecidivism Measures at Follow-Up

Study	Patient population	Alternative treatment	Dependent measures	*N*	*d*	*Z*	Length of follow-up
Foster, Prinz, & O'Leary (1983)	Adolescents with behavior problems.	Education in problem-solving skills.	Self-report questionnaires, experimenter ratings of recorded interactions.	17	−.12	0.0	6–8 weeks
Klein, Alexander, & Parsons (1977)	Adolescents (juvenile offenders).	Individual therapy (client-centered and eclectic/ dynamic).	Number of sibling referrals to the court system.	86	.90	3.89	2.5–3.5 years
Langsley, Machotka, & Flomenhaft (1971) Langsley, Pittman, Machotka, & Flomenhaft (1968)	Adults referred for inpatient treatment.	Hospitalization.	Experimenter ratings of patients, level of functioning.	250	−.09	0.0	18 months
Ro-Trock, Wellisch, & Schoolar (1977); Wellisch & Ro-Trock (1980)	Adolescents and young adults in inpatient treatment.	Individual therapy.	Number of families using outpatient therapy services.	24	−1.22	−1.96	3 years

Discussion

Summary

The results of this review are summarized in Table 8. The results showed that family therapy had a positive effect on patients, compared with no therapy, when the effect was measured by either family interactions or the behavior of the identified patient. For behavior ratings, a large number of studies with null results would have to be uncovered to alter this conclusion. However, only three would reduce the family interaction measures to nonsignificance. The average effect size from both types of outcome was very similar, $d = .45$ and .50, respectively, and the studies were shown to be homogeneous in their result.

When family therapy was compared with alternative types of treatments, the results also demonstrated the relative effectiveness of family therapy. The studies using family interaction measures to compare family therapy with alternative treatments had a larger average effect size than the family therapy versus no-treatment comparison, and this result was statistically significant. However, because of the small number of studies examining this comparison ($n = 3$), only three additional studies with null results would make the overall result nonsignificant. The use of behavior ratings to compare family therapy with alternative treatments showed a smaller positive effect than did family therapy versus no-treatment studies, although the

Table 7
Recidivism Data Comparing Family Therapy and Alternative Treatments

Treatment	Total sample	Hospital patients	Adolescents
Family therapy			
Recidivism	174	69	113
No recidivism	308	117	137
Alternative treatments			
Recidivism	273	120	161
No recidivism	197	130	83
χ^2	46.18	20.93	21.76
p	.005	.005	.005
Effect size (*d*)	.47	.41	.43

Table 8
Family Therapy Compared With No Treatment and Alternative Treatments

Measure	Effect size (*d*)	*Z*	Fail-safe *N*	Homogeneity
		No treatment		
Family interactions	.45	1.88, $p < .03$	3	$\chi(6, N = 7) = 2.94$, $p < .80$
Behavior ratings	.50	3.49, $p < .0002$	26	$\chi(5, N = 6) = 6.57$, $p < .20$
		Alternative treatments		
Family interactions	.65	2.33, $p < .01$	3	$\chi(2, N = 3) = 2.48$, $p < .20$
Behavior ratings	.23	2.63, $p < .004$	7	$\chi(3, N = 4) = 8.10$, $p < .05$
Follow-up (nonrecidivism measures)	.06	0.97, $p = .17$	0	$\chi(3, N = 3) = 23.51$, $p < .005$
Recidivism	.47	2.58, $p < .005$		$\chi(5, N = 512) = 25.73$, $p < .005$

effect was still significant. Again, because of the small number of studies reporting this comparison ($n = 4$), only seven additional studies with null results would make the overall finding nonsignificant. The results of these studies were not homogeneous and the effect sizes ranged from $d = 0.0$ when family therapy was compared with hospitalization to .90 when compared with bibliotherapy. The most obvious possible reason for the heterogeneity is the diverse nature of the alternate treatments. Only one form of alternate treatment appeared in more than one study. Variability in the effectiveness of the alternative treatments would cause differences in the estimate of their relative effectiveness when compared with family therapy.

The results of two studies reporting follow-up data showed that family therapy continued to have positive results compared with no treatment. However, family therapy did not prove to be more effective than alternative treatments when the measure of effectiveness was delayed beyond the time of treatment termination. The effect sizes of these studies were very heterogeneous, probably reflecting the differences in follow-up intervals and dependent measures as well as the use of different alternative treatments. As seen in Table 6, no two studies used the same alternative treatments or dependent variables and only two studies used similar follow-up intervals.

The six studies using recidivism as a follow-up measure of treatment effectiveness did show family therapy to be more effective than alternative treatments. The combined recidivism rates yielded a highly significant result and an effect size comparable to posttreatment measures of effectiveness. However, the separate recidivism rates were shown to be heterogeneous, even among homogeneous populations (i.e., hospital patients or adolescents). Again, this is likely to be due to variations in alternative treatments and the length of follow-up. The correlation between length of follow-up and effect size was negative, showing that there was a trend toward decreased differences between family therapy and alternative treatments with longer follow-up periods.

We also explored the methodological issue of whether study outcomes were affected by random assignment of patients to groups versus patient matching. Little difference in results was associated with this design feature.

The conclusions of this review were based on a variety of dependent measures, client populations, and approaches to family therapy. Because of the small number of studies, the separate effects of many possible mediators could not be statistically evaluated. However, in general, the results of this review are consistently favorable toward family therapy. In addition, positive changes were documented in several areas of family functioning such that family therapy was found to have specific effects on the presenting problem and more general effects on the family system as a whole. This tends to support the contention of pragmatic family therapists that concentrating on behavioral outcomes can provide a vehicle for pervasive systemic change, which is not limited to symptom removal.

Examining Inconsistent Studies and Overall Results

Several studies reported results that were inconsistent with the results of other studies with similar designs. For example, the study by Garrigan and Bambrick (1979) is notable because it found no effect for family therapy compared with no-treatment controls on family interaction variables and a negative effect on ratings of behavior. This was the only study finding null or negative effects when comparing family therapy with no treatment. The reason for this may be that the authors were focusing on differences between single-parent families and intact families in their response to treatment. However, they did report sufficient data to make the more global comparison of all-treated versus all-untreated families. There were numerous differences between single-parent families and intact families and, in fact, the differential responses of these types of families appear to be responsible for the overall lack of positive effects for family therapy. Their data showed that treated single-parent families were more disturbed than the untreated single-parent families following treatment, whereas the treated intact families were less disturbed than the untreated intact families. The authors did not offer any explanation for this difference except to say that the single-parent families were often described by therapists as resistant and unresponsive to treatment.

Another study by the same authors (Garrigan & Bambrick, 1977, 1979) was inconsistent in that it found a much larger effect at follow-up assessment than at posttreatment, whereas most other studies reported smaller effect sizes at follow-up. This study, however, was comparing family therapy with no treatment, instead of to alternative treatments, as did the other studies with follow-up assessments. A plausible explanation for this result is the use of different outcome measures at follow-up (school or employment status and court involvement) than at the posttreatment (ratings of observed family interactions). It is also possible that the no-treatment families continued to worsen and treated families continued to improve, producing an even larger effect at follow-up.

Another curiosity in the findings is illustrated by studies by Langsley et al. (1971, 1968). These studies found differences between family therapy and hospitalization on recidivism rates at follow-up, but other dependent measures did not distinguish the groups at either posttreatment or follow-up. This result was consistent with the general trend in the overall results that follow-up assessment showed virtually no differences on the family interaction and behavior rating measures of family therapy and alternative treatment groups but did show a positive family therapy effect for recidivism, and in some cases the same families were measured in different ways. It is possible that the aspects of family interaction and behavior of the identified patients measured in the studies were not directly related to the source of the families' problems. Family therapy may have had an effect on the families' interactions that was not measured but did reduce recidivism rates for those families. Another possibility is that the family interactions or behaviors that were maintaining the problem were of such low frequency, although with a high impact when they did occur, that any changes were not noted but were sufficient to reduce recidivism.

Another inconsistency among studies was found in the follow-up data from alternative therapy studies using nonrecidivism measures. The effects from these studies were shown to be very heterogeneous (see Table 6): Two studies showed very small effects, and the other two studies had large effects, but in opposite directions. Klein et al. (1977) found that the positive effect of family therapy on siblings' initial arrest rates was actually

larger than on recidivism rates of identified patients. On the other hand, Ro-Trock, Wellisch, and Schoolar (1977) and Wellisch and Ro-Trock (1980) found a negative effect for family therapy. In this study, the alternative treatment (i.e., individual therapy) was administered to various family members instead of only to the identified patient. Thus, this alternate treatment was, in essence, another form of family therapy and may have resulted in changes similar to family therapy with all of the family members present. In addition, the measure used was the number of families that sought treatment at outpatient clinics during a 3-year period following the identified patient's discharge from the hospital. The researchers found that hospitalized patients treated with family therapy were much more likely to seek outpatient treatment during the follow-up period. For the purposes of this review, these results were considered to be negative, on the basis of the assumption that all families that need treatment would seek it out. However, it could also be considered a positive effect for family therapy because of the possibility that families treated with family therapy were more aware of their problems, more motivated to solve them, or viewed treatment positively.

Limitations on the Review's Conclusions

In addition to summarizing the existing literature on family therapy outcome, this review also uncovered several possible moderators of effectiveness that will require more research before assessments of their impact can be made. First, the theoretical position of most family therapists is that family therapy techniques can be used with nearly any client population. However, most of the studies reviewed here studied the effects of family therapy with families of children who showed behavior problems. All of the studies comparing family therapies with no treatment had children's problems as the focus of therapy. In the studies that compared family therapies with alternative treatments, four studies probably included both children and adults as identified patients, although age ranges were not reported (Goldstein et al., 1978; Langsley et al., 1971, 1968; Rittenhouse, 1970; Schubert & Miller, 1981). The effect size for family therapy for these mixed populations did not appear to be different from the effects for children (see Tables 5, 6, and 7), but there are too few studies of similar populations to make a more conclusive statement. Furthermore, within a given population, family therapies may have differential effects depending on the type of disorder being treated (e.g., acting out vs. anxious and withdrawn children). It may also be the case that family therapies have differential effects for nonsymptomatic problems, such as difficulties dealing with transitional phases in the family life cycle. Again, there are too few studies of these hypotheses to draw any conclusions at this time.

Second, family therapies appear to be only slightly more effective than alternative treatments. However, the alternative treatments used in the studies reviewed here were not consistent across the studies. Some studies used alternative treatments that were intended to be legitimate and are in common practice, such as inpatient hospitalization (Langsley et al., 1968). Other studies used groups that received only attention or placebo treatments, such as bibliotherapy (Christensen, Johnson, Phillips, & Glasgow, 1980). It is probable that in our review the results comparing family therapies with alternative treatments are inflated (that is, made more positive toward family therapies) by including these attention/placebo treatment groups. Again, there are too few studies in the literature to reach any definitive conclusions distinguishing between the types of alternate treatments. Before concluding that family therapies are better treatments than other forms of therapy, more and better research must be conducted comparing family therapies with both attention/placebo treatments and viable treatment alternatives.

Finally, a crucial variable in evaluating the effectiveness of family therapies is the type of family therapy interventions used. Family therapies usually include several family members in sessions, but the interventions used in the sessions are not standard or consistent. The family therapy approaches used in the studies reviewed here include structural, strategic, and behavioral family therapies, and parent training in client-centered, Rogerian methods. Furthermore, some studies described the interventions that were used in great detail (e.g., Garrigan & Bambrick, 1975), whereas others reported no information about the type or sequence of interventions that were used (Schubert & Miller, 1981). Because of the lack of complete information about the type and sequence of family therapy interventions used and the extreme variety of family therapy interventions reported, no conclusions can be drawn about differential effects for different types of family therapies.

Future Directions

Clearly, family therapy researchers must strive for greater specificity in the questions they ask, such as "what are the specific effects of specific interventions by specified therapists at specific points in time with particular types of patients with particular presenting problems?" (Gurman et al., 1986, p. 601). However, it may also be the case that there are effective components and change mechanisms that are common to various types of family therapy (Gurman et al., 1986). To the extent that researchers of the family therapy process are able to identify such elements, it may be possible to bridge some of the epistemological rifts that have developed in the field of family therapy.

One fundamental and, as yet, unresolved issue in the evaluation of family therapy outcomes is related to the selection of criteria that are used to define change. Although it is self-evident that there are multiple levels and multiple vantage points for assessing outcomes in family therapies, many of the studies included in this review have taken a narrow approach in this regard. For example, several researchers used parent ratings of patient behavior (Fischer, Anderson, Arveson, & Brown, 1978; Martin, 1977), experimenter ratings of patient behavior (Schubert & Miller, 1981), or recidivism (Johnson, 1977) as the sole criterion for evaluating therapeutic outcome. Undoubtedly, each researcher's values and theoretical biases have a strong impact on both the selection of outcome criteria and on the judgment of what constitutes a positive outcome (Colapinto, 1979; Strupp & Hadley, 1977). Nevertheless, we believe that considerable progress could be made in evaluating family therapies through the increased use of multiple change indices (e.g., individual, marital, and family system variables) from multiple

vantage points (e.g. patient and other family members, therapist, observer). This approach would facilitate between-study comparisons (by ensuring at least some overlap in criteria between studies) and would provide a more detailed representation of the complex effects of family therapies.

Most family therapies are, by definition, designed to effect changes in patterns of intrafamily relations. However, only 8 of the 20 investigations included in this review used measures of family interaction to assess treatment outcomes. We recommend that the use of conceptually relevant family interaction measures become standard practice in family therapy outcome research. However, there are some important methodological issues pertaining to the assessment of family relations. First, considering that intermember agreement about the quality of family relations is generally quite low (see e.g., Alexander, Johnson, & Carter, 1984; Henggeler, Borduin, & Mann, 1987), family therapy researchers should not generalize from the perceptions of one family member to the entire system. Schwarz, Barton-Henry, and Pruzinsky (1985) have recently suggested that aggregate scores based on the ratings of multiple family members provide the most valid description of family relations. However, as Barnes and Olson (1985) have noted, it may also be useful to examine intrafamily differences on variables of interest. To the extent that family members differ in their perceptions of family relations, description of these differences may advance our understanding about the nature of family functioning and family change.

Second, although expensive and time consuming, increased use of observational methods in studies of family therapy outcome is clearly needed. Observational measures permit sequential analyses of family interaction data and assess bidirectional influences within the family system more adequately than do self-report measures. However, we do not wish to imply that observational measures provide a more accurate picture of family relations than do self-report measures. Because self-report and observational methods provide different vantage points for viewing current family interactions, family therapy researchers should use both types of methods whenever possible.

Although the family is certainly one of the most important systems that influence the individal, there is little doubt that the individual is also embedded in systems that extend beyond the family (Bronfenbrenner, 1979; Henggeler & Borduin, in press; Salzinger, Antrobus, & Glick, 1980). For example, the older child or adolescent often has extensive transactions with peers, school personnel, and neighborhood residents. For the adult, primary extrafamilial systems typically include the work environment, friends, extended family, and other social support networks. From a multisystemic perspective, therapeutic efforts to alter directly interactional exchanges between family members may have indirect effects on the many interrelated systems in the family's ecological environment. As a consideration of extrafamilial systems seems essential for a more complete understanding of the various factors that are related to psychosocial difficulties, future investigations of family therapy outcomes should broaden their focus to include measures of relevant extrafamilial variables.

There are several other relevant outcome criteria that deserve consideration in studies of family therapy effectiveness. Alternatives to between-group comparisons of average patient/family performance include evaluations of the importance (i.e., clinical significance) of family change, the proportion of families who improve, and the durability of improvement (Jacobson, Follette, & Revenstorf, 1984; Kazdin & Wilson, 1978). In addition, factors such as the duration of therapy, the costs associated with professional training, and the cost and acceptability of therapy to the patient or family might all represent relevant criteria for comparing different treatments. In fact, the rise in third-party payments may soon demand evidence of treatment efficiency and consumer satisfaction regarding various forms of treatment, including family therapies. On the other hand, as consumer satisfaction may not be equivalent to effective treatment (Garfield, 1983; Parloff, 1983), researchers and policymakers must exercise caution in evaluating consumer data. Obviously, such evidence could be used in conjunction with other types of outcome data and might be weighted accordingly.

Conclusion

Until an extensive set of family therapy outcome studies has accumulated, any assessment of the effectiveness of family therapies must remain tentative. This review indicates that family therapies do work, in their general form, and seem to work better than some alternative treatments. Although these conclusions are tentative, they should be encouraging to therapists in this young, developing field. Ironically, as researchers continue to establish an empirical base regarding the efficacy and outcomes of family therapy, they will need to do so in the face of challenges from advocates of the new epistemology (see, e.g., Colapinto, 1979; Tomm, 1983). These theorists have charged, among other things, that traditional research methods are derived from linear and reductionistic paradigms and, therefore, are inappropriate and inadequate to contribute to our knowledge of how systems operate and change. However, as Gurman (1983) has cogently argued, standard research methods are the only ethically responsible means currently available for evaluating family therapy outcomes. Moreover, researchers of family therapy outcome and process can attend to context by studying the interactions of patient, therapist, setting, and treatment variables and by using multiple levels and vantage points for assessing systemic change. We hope the results of this review will encourage researchers to produce better studies evaluating the multidimensional effects of family therapies and involving variations in patients, outcomes, and alternative treatments.

References

Alexander, B. B., Johnson, S. B., & Carter, R. L. (1984). A psychometric study of the Family Adaptability and Cohesion Evaluation Scales. *Journal of Abnormal Child Psychology, 12,* 199–208.

Alexander, J., & Parsons, B. (1973). Short-term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 81,* 219–225.

Attneave, C. L. (1976). Social networks as the unit of intervention. In P. J. Guerin, Jr. (Ed.), *Family therapy: Theory and practice* (pp. 220–232). New York: Gardner.

Barnes, H. L., & Olson, D. H. (1985). Parent–adolescent communication and the circumplex model. *Child Development, 56,* 438–447.

Bogart, A., & French, A. (1978, February). Successful short-term fam-

ily therapy with incarcerated adolescents. *Journal of Juvenile and Family Courts,* 3–8.

Borduin, C., Henggeler, S., Hanson, C., & Harbin, F. (1982). Treating the family of the adolescent: A review of the empirical literature. In S. W. Henggeler (Ed.), *Delinquency and adolescent psychopathology: A family–ecological systems approach* (pp. 205–222). Boston: John Wright–PSG.

Bowen, M. (1978). *Family therapy in clinical practice.* New York: Aronson.

Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge, MA: Harvard University Press.

Christensen, A., Johnson, S., Phillips, S., & Glasgow, R. (1980). Cost effectiveness in behavioral family therapy. *Behavior Therapy, 11,* 208–226.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Colapinto, J. (1979). The relative value of empirical evidence. *Family Process, 18,* 427–441.

Cook, D., & Campbell, D. (1979). *Quasi-experimentation.* Chicago: Rand McNally.

Cooper, H. (1979). Statistically combining independent studies: A meta-analysis of sex differences in conformity research. *Journal of Personality and Social Psychology, 37,* 131–146.

Cooper, H. (1984). *The integrative research review: A systematic approach.* Beverly Hills, CA: Sage.

Dell, P. H. (1982). Beyond homeostasis: Toward a concept of coherence. *Family Process, 21,* 21–41.

DeWitt, K. (1978). The effectiveness of family therapy: A review of outcome research. *Archives of General Psychiatry, 35,* 549–561.

Douds, A., Engelsgjerd, M., & Collingwood, T. (1977). Behavior contracting with youthful offenders and their parents. *Child Welfare, 56,* 409–413.

Fisch, R., Weakland, J. H., & Segal, L. (1982). *The tactics of change: Doing therapy briefly.* San Francisco: Jossey-Bass.

Fischer, J., Anderson, J., Arveson, E., & Brown, S. (1978). Alderian family counseling: An evaluation. *International Journal of Family Counseling, 6,* 42–44.

Foster, S., Prinz, R., & O'Leary, K. (1983). Impact of problem-solving communication training and generalization procedures on family conflict. *Child and Family Behavior Therapy, 5,* 1–23.

Friedman, H. (1968). Magnitude of experimental effect and a table for its rapid estimation. *Psychological Bulletin, 70,* 245–251.

Garfield, S. L. (1983). Some comments on consumer satisfaction in behavior therapy. *Behavior Therapy, 14,* 237–241.

Garrigan, J., & Bambrick, A. (1975). Short term family therapy with emotionally disturbed children. *Journal of Marriage and Family Counseling, 1,* 379–343.

Garrigan, J., & Bambrick, A. (1977). Family therapy for disturbed children: Some experimental results in special education. *Journal of Marriage and Family Counseling, 3,* 83–93.

Garrigan, J., & Bambrick, A. (1979). New findings in research on go-between process. *International Journal of Family Therapy, 1,* 76–85.

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

Goldenberg, I., & Goldenberg, H. (1985). *Family therapy: An overview* (2nd ed.). Monterey, CA: Brooks/Cole.

Goldstein, M., Rodnick, E., Evans, J., May, P., & Steinberg, M. (1978). Drug and family therapy in the aftercare of acute schizophrenics. *Archives of General Psychiatry, 35,* 1169–1177.

Gould, E., & Glick, I. (1977). The effects of family presence and brief family intervention on global outcome for hospitalized schizophrenic patients. *Family Process, 16,* 503–510.

Gruher, M. (1979, February). Family counseling and the status offender. *Juvenile and Family Court Journal,* 23–27.

Gurman, A. S. (1983). Family therapy research and the "new epistemology." *Journal of Marital and Family Therapy, 9,* 227–234.

Gurman, A. S., & Kniskern, D. P. (1978). Research on marital and family therapy: Progress, perspective, and prospect. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (pp. 817–902). New York: Wiley.

Gurman, A. S., Kniskern, D. P., & Pinsof, W. M. (1986). Research on the process and outcome of marital and family therapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (pp. 565–624). New York: Wiley.

Haley, J. (1976). *Problem-solving therapy.* New York: Harper & Row.

Hardcastle, D. (1977). A mother–child, multiple-family counseling program: Procedures and results. *Family Process, 16,* 67–74.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hendricks, W. (1971). Use of multifamily counseling groups in treatment of male narcotic addicts. *International Journal of Group Psychotherapy, 21,* 84–90.

Henggeler, S. W. (Ed.). (1982). *Delinquency and adolescent psychopathology: A family-ecological systems approach.* Boston: John Wright–PSG.

Henggeler, S. W., & Borduin, C. M. (in press). *Beyond family therapy: A multisystemic approach to treating the behavior problems of children and adolescents.* Homewood, IL: Dorsey Press.

Henggeler, S. W., Borduin, C. M., & Mann, B. J. (1987). Intrafamily agreement: Association with clinical status, social desirability, and observational ratings. *Journal of Applied Developmental Psychology, 8,* 97–111.

Hetherington, E., Stouwie, R., & Ridberg, E. (1971). Patterns of family interaction and child-rearing attitudes related to three dimensions of juvenile delinquency. *Journal of Abnormal Psychology, 78,* 160–176.

Hoffman, L. (1981). *Foundations of family therapy.* New York: Basic Books.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15,* 336–352.

Johnson, T. (1977). The results of family therapy with juvenile offenders. *Juvenile Justice, 23,* 29–33.

Johnson, T., & Maloney, H. (1977). Effects of short-term family therapy on patterns of verbal interchange in disturbed families. *Family Therapy, 4,* 207–213.

Katz, A., de Krasinski, M., Philip, E., & Wieser, C. (1975). Change in interactions as a measure of effectiveness in short-term family therapy. *Family Therapy, 2,* 31–56.

Kazdin, A. E., & Wilson, G. T. (1978). Criteria for evaluating psychotherapy. *Archives of General Psychiatry, 35,* 407–418.

Keeney, B. P. (1983). *The aesthetics of change.* New York: Guilford.

Keeney, B. P., & Sprenkle, D. H. (1982). Ecosystemic epistemology: Critical implications for the aesthetics and pragmatics of family therapy. *Family Process, 21,* 1–19.

Klein, N., Alexander, J., & Parsons, B. (1977). Impact of family systems intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. *Journal of Consulting and Clinical Psychology, 45,* 469–474.

Langsley, D., Machotka, P., & Flomenhaft, K. (1971). Avoiding mental hospital admissions. *American Journal of Psychiatry, 127,* 1391–1394.

Langsley, D., Pittman, F., Machotka, P., & Flomenhaft, K. (1968). Family crisis therapy—results and implications. *Family Process, 7,* 145–158.

Martin, B. (1977). Brief family intervention: Effectiveness and the importance of including the father. *Journal of Consulting and Clinical Psychology, 45,* 1002–1010.

Massie, H., & Beels, C. (1972). The outcome of the family treatment of schizophrenia. *Schizophrenia Bulletin, 6,* 24–36.

Masten, A. (1979). Family therapy as a treatment for children: A critical review of outcome research. *Family Process, 18,* 323–335.

Minuchin, S. (1974). *Families and family therapy.* Cambridge, MA: Harvard University Press.

Montalvo, B., & Haley, J. (1973). In defense of child therapy. *Family Process, 12,* 227–244.

Nichols, M. P. (1984). *Family therapy: Concepts and methods.* New York: Gardner.

O'Connor, W. A., & Lubin, B. (Eds.). (1984). *Ecological approaches to clinical and community psychology.* New York: Wiley.

Olson, D., Russell, C., & Sprenkle, D. (1979). Circumplex model of marital and family systems: I. Cohesion and adaptability dimensions, family types, and clinical applications. *Family Process, 18,* 3–28.

Ostensen, K. (1981). The runaway crisis: Is family therapy the answer? *The American Journal of Family Therapy, 9,* 3–12.

Paquin, M. (1977). The status of family and marital therapy outcomes: Methodological and substantive considerations. *Canadian Psychological Review, 18,* 221–232.

Parloff, M. B. (1983). Who will be satisfied by "consumer satisfaction" evidence? *Behavior Therapy, 14,* 242–246.

Parsons, B., & Alexander, J. (1973). Short-term family intervention: A therapy outcome study. *Journal of Consulting and Clinical Psychology, 48,* 195–201.

Patterson, G. (1971). *Families: Applications of social learning to family life.* Champaign, IL: Research Press.

Patterson, G. R. (1982). *Coercive family process.* Eugene, OR: Castalia.

Pevsner, R. (1982). Group parent training versus individual family therapy: An outcome study. *Journal of Behavior Therapy and Experimental Psychiatry, 13,* 119–122.

Reiter, G., & Kilmann, P. (1975). Mothers as family change agents. *Journal of Counseling Psychology, 22,* 61–65.

Rittenhouse, J. (1970). Endurance of effect: Family unit treatment compared to identified patient treatment. *Proceedings of the 78th Annual Convention of the American Psychological Association, 5,* 535–536.

Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin, 85,* 185–193.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Rosenthal, R. (1984). *Meta-analysis procedures for social research.* Beverly Hills, CA: Sage.

Ro-Trock, G., Wellisch, D., & Schoolar, J. (1977). A family therapy outcome study in an inpatient setting. *American Journal of Orthopsychiatry, 47,* 514–522.

Russell, C., Olson, D., Sprenkle, D., & Atilano, R. (1983). From family symptom to family system: Review of family therapy research. *American Journal of Family Therapy, 11,* 3–14.

Salzinger, S., Antrobus, J., & Glick, J. (Eds.). (1980). *The ecosystem of the "sick" child.* New York: Academic Press.

Schubert, D., & Miller, S. (1981). Social class and perceived improvement in therapy: The effects of therapist discipline and therapy type. *The Hillside Journal of Clinical Psychiatry, 3,* 71–80.

Schwarz, J. C., Barton-Henry, M. L., & Pruzinsky, T. (1985). Assessing child-rearing behaviors: A comparison of ratings made by mother, father, child, and sibling on the CRPBI. *Child Development, 56,* 462–479.

Sigal, J., Barrs, C., & Doubilet, A. (1976). Problems in measuring the success of family therapy in a common clinical setting: Impasse and solutions. *Family Process, 15,* 225–233.

Slipp, S., & Kressel, K. (1979). Does family therapy produce change? An overview of outcome studies. *Group, 3,* 23–34.

Stover, L., & Guerney, B. (1967). The efficacy of training procedures for mothers in filial therapy. *Psychotherapy: Theory, Research and Practice, 4,* 110–115.

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcomes: With special reference to negative effects in psychotherapy. *American Psychologist, 32,* 187–196.

Stuart, R., Jayaratne, S., & Tripodi, T. (1976). Changing adolescent deviant behavior through reprogramming the behavior of parents and teachers: An experimental evaluation. *Canadian Journal of Behavioral Science, 8,* 132–144.

Szapocznik, J., Kurtines, W., Foote, F., Perez-Vidal, A., & Hervis, O. (1983). Conjoint versus one-person family therapy: Some evidence for the effectiveness of conducting family therapy through one person. *Journal of Consulting and Clinical Psychology, 51,* 889–899.

Tomm, K. (1983). The old hat doesn't fit. *Family Therapy Networker, 7,* 39–41.

Ulrici, D. (1983). The effects of behavioral and family interventions on juvenile recidivism. *Family Therapy, 10,* 25–36.

Weathers, L., & Liberman, R. (1975). Contingency contracting with families of delinquent adolescents. *Behavior Therapy, 6,* 356–366.

Wellisch, D., & Ro-Trock, G. (1980). A three-year follow-up of family therapy. *International Journal of Family Therapy, 2,* 169–175.

Wells, R., & Dezen, A. (1978). The results of family therapy revisited: The nonbehavioral methods. *Family Process, 17,* 251–273.

Whitaker, C. (1976). A family is a four-dimensional relationship. In P. J. Guerin, Jr. (Ed.), *Family therapy: Theory and practice* (pp. 182–192). New York: Gardner.

Wiltz, N., & Patterson, G. (1974). An evaluation of parent training procedures designed to alter inappropriate aggressive behavior of boys. *Behavior Therapy, 5,* 215–221.

Received May 2, 1986
Revision received October 20, 1986 ■