

ON THE RELATIONSHIP BETWEEN THERAPIST EXPERIENCE AND PSYCHOTHERAPY OUTCOME

David M. Stein

Indiana State University

and

Michael J. Lambert

Brigham Young University

ABSTRACT. *The question of whether a relationship exists between the therapists' level of experience or training and psychotherapy outcome has numerous implications for clinical training and practice. A comprehensive examination and meta-analysis of available research suggests that answers to this question are complex and problematic. The present review considers the question of the relationship between experience and outcome as it relates to a number of therapist, client, treatment, and methodological variables. This paper also considers past attempts by reviewers to summarize studies that sought to address this issue. In addition, it examines variables that helped predict whether a study's results favored (to varying degrees) experienced clinicians versus less-trained persons. A discussion of prominent methodological problems and recommendations for future research is also offered.*

It has long been assumed that formal, professional training in psychotherapy is a necessary prerequisite to providing quality mental health services. Indeed, current systems of graduate-level training in the helping professions are based largely on this assumption. However, a number of investigators have begun to challenge this presumption, noting that it may have been largely reinforced by clinical impressions. Strupp, Fox, and Lessler (1969), for example, have suggested that the enthusiasm of the novice psychotherapist may compensate for the lack of technical skill he/she may possess. Gomes-Schwartz (1978) has emphasized that the outcomes of psychotherapy as based on current measures may show less variability across level of training and experience than might have been previously assumed. Still other writers have questioned the need for any formal, professional training. One

Requests for reprints should be addressed to David M. Stein, Department of Psychology, Indiana State University, Terre Haute, IN, 47809.

extreme view is that therapists who are products of professional programs often do clients more harm than good and that lay helpers are more therapeutic (e.g., Carkhuff, 1968).

It is becoming increasingly important to clearly outline the empirical evidence that bears on the experience–outcome issue. Such evidence would have implications for current graduate training models in psychotherapy. The experience–outcome question also has import for the recent, rapid growth of lay and paraprofessional training programs. A basic concern associated with nontraditional training programs and the use of lay therapists is the question of who is qualified to provide direct services to the public. Additionally, factors that might qualify conclusions regarding the therapists' level of training or experience, and their relationship to client change deserve careful consideration.

OVERVIEW OF PAST REVIEWS

Several prominent reviewers have attempted to uncover patterns in outcome across studies dealing with therapists' level of training and experience using various methods of organization. Unfortunately, there has been little agreement regarding the conclusions to be drawn from this literature. This failure is probably due to several factors, such as the use of different samples of studies and diverse review methods.

One type of review strategy has been to examine samples of *general psychotherapy outcome studies*, attempting to uncover patterns of effectiveness as they may be related to therapist experience. Bergin (1971) coded the level of therapist experience across 48 studies of general psychotherapy outcome. He reported that 53% of his sample of studies suggested positive results for experienced therapists, while only 18% of the studies that used inexperienced therapists showed clear improvement.

Using a similar between-studies approach, Smith and Glass (1977) correlated therapists' years of experience with standardized outcome measures from 475 studies that investigated the general effectiveness of psychotherapy. Contrary to Bergin's conclusions, Smith and Glass found no relationship between years of experience and therapy outcome ($r = .00$). They caution that therapists in typical psychotherapy studies tend to be quite inexperienced (e.g., only about 30% had 5 or more composite years of training and experience).

As part of their examination of the relative efficacy of different treatment approaches, Shapiro and Shapiro (1982) correlated estimates of therapist years of experience with size of treatment effect across studies. The average therapist in these studies was even younger than in the aforementioned Smith et al. review (mean 2.91 years of *training plus experience*). Like Smith et al., no overall relationship between experience and outcome was found when all patient target problems were considered. However, there was a tendency across studies for novice therapists to treat circumscribed behavior problems that were highly reactive and which produced larger treatment effect sizes.

The fundamental problem with reviews that assess the experience–outcome relationship *across studies* is the restricted range of experience of therapists typically used in psychotherapy outcome studies. Also, a basic confidence must exist in the validity of methods used to categorize estimates of level of training and experience of therapists. Both Smith et al. and Shapiro and Shapiro experienced some difficulty here. Glass and Smith (1980) outline a final problem bearing on the interpretation

of results using this approach. They emphasize that one has no assurance that experienced and less-experienced psychotherapists in the diverse studies they examined worked with comparable clients in similar treatment settings.

A second strategy for reviewing studies that address the question of whether therapist experience relates to outcome has been to directly compare the effectiveness of therapists of differing levels of experience *within a study*. In what is probably the weakest review of studies of this type to date, Anthony and Carkhuff (1977) assessed an unrepresentative and small sample of studies pertaining to the effectiveness of lay therapists, relative to experienced clinicians. Their central conclusion was that regardless of the level of training, experience, or supervision, lay therapists are as effective as the professional clinician.

In addition, Durlak (1979) recently evaluated over 40 studies that directly compared "paraprofessional" and "professional" therapists. Paraprofessionals were therapists who had not received traditional, graduate-level training in psychotherapy. Often, they were college students and other lay helpers. The professionals, on the other hand, had generally completed graduate-level training and had varying levels of experience. Durlak concluded that, regardless of the quality of the research design used, the psychotherapy outcomes of paraprofessionals were as good if not better than those of professional mental health workers.

The two aforementioned reviews by Anthony and Carkhuff (1977) and Durlak (1979) are problematic from several perspectives. Anthony and Carkhuff used an unrepresentative sample of studies and failed to examine outcomes in relation to internal validity variables. Durlak's review of research pertaining to the relative effectiveness of paraprofessionals has been critiqued adequately by Fisher and Nietzel (1980). Overall, Fisher and Nietzel reject nearly all of the studies Durlak included in his review. Like Anthony and Carkhuff (1977), Durlak drew conclusions regarding the effectiveness of lay therapists from studies that were in fact, largely inconclusive, ambiguous and had serious internal validity problems. The two reviews by Anthony and Carkhuff (1977) and Durlak (1979) also focus on a restricted subgroup of inexperienced therapists. The so-called paraprofessional or lay therapist may not be representative of other less-trained personnel, such as graduate students enrolled in professional mental health programs.

Another recent but more carefully considered review conducted by Auerbach and Johnson (1977) examined a narrow but different subset of comparative outcome studies than was used by Durlak, or Anthony and Carkhuff. Basically, they viewed within-study comparisons of *professional trainees and novice clinicians* relative to more senior psychotherapists (while omitting studies of paraprofessionals). Auerbach and Johnson concluded that a modest relationship between experience and outcome does exist, although it proved to be weaker than they expected. Also, experienced clinicians seem to form better therapeutic relationships with client than do novice.

SUMMARY AND APPRAISAL

The available reviews have examined the question of whether therapist experience is related to client improvement from diverse but restricted perspectives. Reviews have included studies involving different clinical and therapist populations and

have used a variety of strategies for summarizing the literature. Overall, the trends in the conclusions of past reviews are somewhat mixed but tend to discount experience as a relevant variable. Within-study comparisons of outcomes for therapists of different levels of experience hold the greatest potential for controlling confounding variables. However, past reviews of such studies have used diverse samples of studies and problematic methods for summarizing and integrating results. These problems suggest that a more comprehensive and less problematic review of the literature on this topic may be useful.

The present review attempts to provide a broader examination of therapist experience as it relates to psychotherapy outcome using improved methods (e.g., meta-analysis) for integrating studies. Several pertinent questions will be points of focus. Is there a general relationship between experience and outcome? How do the definitions of "experience" used in the current literature and different types of therapists relate to sizes of treatment effect? Do various sources or perspectives on assessing treatment effectiveness relate to the size of treatment effect for experienced and less-experienced therapists? Finally, are there any general variables that help predict whether a study produced results favoring less-trained therapists, versus those that support the traditional notion of the advantage of experience?

METHOD

Selection of Studies for Review¹

The primary focus of the literature search for the present review was to assess studies dealing with individual or group psychotherapy. It was considered important to select studies targeted on real clinical problems using such treatment approaches as psychodynamic, client-centered therapy, and behavioral methods (e.g., desensitization). Authors' descriptions of therapy modalities were taken at face value as detailed; operational definitions of treatment were rarely offered.

About 60 studies were obtained from past reviews of the literature on this topic and from computer and hand searches of *Psychological Abstracts*. However, not all of these studies met inclusion criteria. For example, studies involving treatment of drug abuse, neurological problems, educational remediation, and vocational guidance were omitted from consideration. Also rejected were studies examining the use of psychiatric technicians as so-called "primary care therapists" with inpatients, and the use of parents as behavior change agents for their own children. Analogue studies entailing use of pseudo-clients, mock interviews and clinical problems not deemed to be clinically relevant or typical (e.g., snake phobias) were also not used. Finally, studies were omitted where fewer than three therapists comprised either the experienced or less-experienced therapist group, or where training or experience differences between groups were impossible to judge. These last two criteria caused the omission of numerous investigations of behavioral techniques.

The final listing of relevant studies meeting aforementioned inclusion criteria is provided in Table 1. The studies are categorized according to source/type of outcome and concomitant sizes of treatment effect.

¹A table of these studies including relevant client, therapist, treatment and methodology variables is available upon request from the first author. The table also contains some widely cited studies not meeting the inclusion criteria used in the present review.

TABLE 1. Source of Outcome and Psychotherapy Effects

Author	Outcome Type	Mean Effect Size
A. Therapist Sources		
Barrett-Lennard (1962)	Therapist Global Assessment	-.71
Bloom & Trautt (1978)	Therapist Global Adjustment	-.68
Brown (1970)	Mean of Scales on Perception & Evaluation of Treatment	+.65
Cartwright et al. (1963)	Therapist Composite Score	.00
Feifel & Eells (1963)	Therapist Rated Overall Adjustment	.00
Hill (1975)	Therapist Global Satisfaction	+.27
Karlsruher (1974)	Therapist Global Rating	+2.77 ^a
Kirshner et al. (1978)	Mean Global Ratings	-.22
Lerner (1972)	Therapist Global Ratings	-.33
Luborsky et al. (1980)	Composite of Client Gain Indices	.00
Messer & Boals (1981)	Mean of Therapist Global Assessments	.00
Scher (1975)	Mean of Therapist Global Assessments	.00
Strupp & Hadley (1979)	Mean of all Therapist Ratings	-.23
Sullivan et al. (1958)	Mean of Therapist Global Assessments	.00
Tuma et al. (1978)	Composite of Therapist Scale Ratings	.00
B. Client Sources		
Beck et al. (1981)	Client Overall Adjustment	-.83
Bloom & Trautt (1978)	Patient Global Adjustment	+.25
Cartwright et al. (1963)	Composite Global Change in Self-Eval. MMPI	.00
Cartwright & Vogel (1960)	Q-Adjustment Scores	-.40
Cole et al. (1969)	Mean of Self-Report Indices	.00
Feifel & Eells (1963)	Patient-rated Overall Adjustment	.00
Grigg (1961)	Client Global Assessment	.00
Karlsruher (1974)	California Test of Personality	+1.06
Kirshner et al. (1978)	Patient Mean Global Ratings	-.20
Lerner (1972)	Rorschach-derived Scores	.00
	Client Outcome Ratings	-.10
Luborsky et al. (1980)	Composite of Gain Indices for Clients	.00
Messer & Boals (1981)	TAT-derived Scores	-.10
Poser (1966)	Mean of a Variety of Perceptual-Motor & Performance Scores	+.27
Russell & Wise (1976)	Mean of Sum of all Self-Report Scales	+.27
Scher (1975)	Mean of 5 indices dealing with symptoms and satisfaction	-.35
Shelton & Madrazo-Peterson (1978)	Suinn Test Anxiety & Symptom Checklist	-.01
	Grade Point Average	-.12
Strupp & Hadley (1979)	Mean of Client Source Indices	+.30
Tuma et al. (1978)	MMPI index	.00
C. Observer/Judge Sources		
Cartwright et al. (1960)	Diagnosticians Global Judgement	.00
Karlsruher (1974)	Briston Social Adjustment Scale	+.04
Levitz & Stunkard (1974)	Specific behavior monitoring	-.95
Luborsky et al. (1980)	Composite of Client Gain Indices.	.00
O'Brien et al. (1972)	Experimenter's Composite Score on Outcome	.00
Strupp & Hadley (1979)	Mean of Independent Clinician's Scores	-.19
Tuma et al. (1978)	Composite of Observers Scales	.00

^aThis score was considered an "outlier" in discriminant analysis.

Review Procedures and Criteria

Two raters independently coded studies for presence or absence of certain client, therapist, treatment, and methodology variables. This coding procedure paralleled one used by Smith and Glass (1977) in their meta-analysis of studies of psychotherapy. Overall, inter-rater agreement on judgments for these categories was quite good, never dropping below 85% for any variable. Two raters independently coded studies. Some of the variables of primary interest included, severity of disorder, type and duration of treatment, number of therapists in each group, and predominant school of training. Other independent variables used to address the question at hand centered on research design and outcome measure issues. Studies were coded for the presence or absence of untreated and placebo control groups, the number of clients seen, how clients were assigned to therapists, attrition rates, standardized length and similarity of treatment across groups, and reliability and blindness of raters.

Attempts were made to examine size of treatment effect as a function of a rough categorization of the *type* of treatment, as well as location (e.g., school, hospital, community mental health center). Thus, outcome measures were categorized according to *source* and *type*. Source refers to who completed the assessment (i.e., client, therapist, patient, judge), while "type" categories refer to the object of assessment (e.g., fear/anxiety, self-esteem, global rating of adjustment, social skills, personality traits, and physiological measures).

Definition of Experience and Related Variables. Of particular concern in the present review was the accurate coding of therapist experience. It was found that "therapist experience" was usually defined in studies as the number of years of applied training (years involving practicum internship, or residency), *plus* post-degree years of clinical practice. Thus, training and experience are typically confounded in available studies. In addition, therapist age is not controlled for in the studies and is confounded with level of experience. In an attempt to deal with the relatively crude and problematic indices of experience commonly found, the authors of studies were contacted directly whenever possible to more clearly ascertain the level of experience of therapists. It was hoped that this would improve the validity of estimates of therapists' experience.

Three similar approaches to defining and examining therapist experience were used in the present review. In the first, therapists were assigned a score based on years of training *plus* experience: M.D., Ph.D. level clinicians were assigned a mean score based on 4 years, plus the group mean estimate of post-degree years of clinical experience. Scores for psychiatric nurses and other master's level clinicians were based on 2 years, plus post-degree experience. For trainees and paraprofessionals, length of training (or lack of it) was estimated similarly. For studies using highly experienced clinicians, this approach provided scores weighted more heavily on clinical experience than formal training. The mean years of experience was coded for each group of therapists in a study.

The other approach used an *experience difference score* (years of experience for the experienced group, minus that of the less-experienced group). This provided the means for examining the general distance between therapist groups being compared on the experience dimension.

A third strategy for examining experience involved the coding of a variable

relating more to the school of training of the therapists than years of clinical experience. This coding system for deriving the variable is as follows: A = college students, lay persons, medical student, school teacher etc., with minimal or no training; B = novice graduate students in mental health professions; C = trained graduate students (1–2 yrs of practicum or equivalent); D = minimally experienced degree recipients: MS, Ph.D., MD, with 1–2 years post-degree experience; E = experienced degreed clinicians with 3 or more years of post-degree experience. A combination variable was created for each study, reflecting the training/experience for the therapist groups being compared in a study e.g., “AC”, “CE” etc.

Size of Effect. Studies of therapy effectiveness used a diversity of outcome measures, making comparisons of results difficult. A central problem of past reviews is that they have used unstandardized methods for comparing diverse outcome measures across studies. It is likely that subjective judgments regarding conclusions to be drawn have resulted. Smith and Glass (1977) suggest that the *size of effect* may provide a superior method for interpreting the results of studies. The size of effect is a transformed, standardized score of outcome that allows for direct comparison of improvement across outcome measures. For the present review, the experienced therapist group is used as the reference or “control” against which the performance of less-trained therapists is compared. The size of effect is calculated:

$$\frac{\bar{x}_{\text{inexper.}} - \bar{x}_{\text{exper.}}}{s_{\text{exper.}}}$$

The mean of the posttest score for the experienced group is subtracted from the less-experienced therapist group. This difference is divided by the standard deviation of the experienced group. Conventions for deriving effect sizes when other summary statistics are provided by the author (*F*-ratios, *t*-tests, significance levels, etc.) are readily available (cf. Smith et al., 1980). A size of treatment effect that is negative reflects support for the traditional assumption that inexperienced therapists are to some degree, associated with poorer outcomes relative to experienced clinicians.

Some studies assessed client change from numerous perspectives (sources). To avoid undue weighting of overall mean effect sizes from studies that used multiple outcome measures, the rule was generally followed that a study could contribute only *one mean effect size per source*. This rule was violated on a few occasions when *types* of outcome measures were especially distinctive or unique for a giver source (e.g., personality tests vs. global self-report measures). Thus, one additional measure was allowed for rare occasions of unique *types* of outcome within a source category. This convention did not affect the overall outcome of the review. Results were analyzed with and without exceptions to the rule.

Also, the data to be presented were based on the convention that a study could contribute up to three measures, one for each of the possible *sources*: client, therapist, or judge. However, on the average, a study contributed two measures to the data pool. The present review involved 46 effect sizes. It should be kept in mind that since studies provided more than one measure, the 46 indices are not independent.

Data Analysis

The major analyses involved examinations of frequency distributions and descriptive statistics of study characteristics, multiple regression analysis, and discriminant analysis.

In the regression, a combination of stepwise and standard multiple regression analysis was used to eliminate categories of independent variables that did not predict effect size (the dependent variable). Group means were substituted for occasional missing data points where necessary. Squared, semi-partial correlations for the final significant variables in the regression equation highlighted the unique contribution of each variable toward explaining the variance in effect size.

The basic discriminant analysis sought to uncover variables that predicted a study's membership into one of two groups: (1) studies that support the traditional notion regarding the importance of experience and outcome; and (2) studies publishing findings contrary to this presumption.

RESULTS

Characteristics of Typical Studies

Based on frequency distributions of the independent variables, the following trends are typically found in the literature to date.

Client Characteristics. The typical client in these comparative studies was coded as having a clinically significant problem (e.g., "neurosis" or personality disorder of mild to moderate severity). Seventy percent of clients in studies scored at the midpoint of a 5 point scale of our "severity" variable. Equal numbers of patients fell on either side of this mean. Clients were almost always self-referred for treatment, rather than solicited or recruited by the experimenter.

Treatment Variables. Therapy in these studies was generally outlined in nonspecific terms. In almost half of the studies, therapists were either free to use any treatment they deemed appropriate, or "therapy" was construed in a very vague manner. Across studies, about 20% of patients were said to receive client-centered therapy. The rest of the clients were evenly distributed across nonspecific counseling, psychodynamic, and behavioral modalities. The median length of therapy in the studies was 20 sessions; however, studies varied greatly on this dimension, ranging from 5 to 125 hours.

In 40% of studies, college counseling centers served as the locations for treatment; some of these centers served only students. Other settings (e.g., mental health center schools) were about equally represented, with the exception of such settings as private hospitals and clinics in residential settings (which were never used).

Therapists. Studies usually included about 9 therapists in each of the comparison groups. The therapist in the "experienced" group usually had between 6 and 7 years of training-plus-experience. These therapists had about 2–3 years more

clinical experience (on the average) than therapists mentioned in the psychotherapy outcome literature generally (cf. Shapiro & Shapiro, 1982; Smith & Glass, 1977). The inexperienced group averaged 1.4 years of accumulated training and experience. Psychology and counseling were coded as schools of training in 30–40% of all studies. “Paraprofessional,” social work, and psychiatric schools of training were equally represented among the remainder of the studies.

Methodology of Typical Studies. Almost all studies failed to include a third, no-treatment control group. Thus, effect sizes of zero are probably best interpreted as inconclusive or ambiguous. The majority of studies did not standardize the length of treatment, or carefully monitor the number of clients seen by the therapists in each group. Almost without exception, clients were assigned on a rotating basis, making the assumption of randomization questionable. Attrition rates were impossible to assess in most cases. In some instances, therapists merely picked up new clients, or failed to report overall caseload figures. Overall, it was difficult to assess how comparable treatment groups were in most studies. However, we feel that this state of affairs is quite typical in the outcome research literature published to date. As has been outlined, methodological problems were coded as independent variables to be tested as predictors of effect size or outcome.

How Researchers Measured Change. The majority of studies included *client assessments* of outcome as a primary evaluation *source*. One-third of studies relied on therapists, either singly or in combination with other sources to assess outcome. Assessments were made by observers or judges only infrequently; these assessments were about as likely as not to be blind and reliable.

One type of outcome category did predominate. About half of the time, *global assessments of adjustment, outcome, or satisfaction* were used by researchers. Personality trait measures were used 13% of the time, with the 9 other *types* of outcome providing the remainder of the evaluations.

Size of Effect. What is the average size of effect across studies? Of 46 outcome measures transformed to effect sizes, about half were 0. Of the remainder, 17 ranged between approximately -1.9 and 0. The rest were positive, (reflecting somewhat more favorable, relative outcomes for novices). The overall mean effect size was 0 (*sd.* = .53).

Hypotheses of Present Review

A series of stepwise and standard regression analyses were run on sets of variables to uncover any significant predictors of effect size. However, the data showed that none of the hypothesized relationships between study variables *thought to be conceptually related to effect size* were found to be useful predictors. These included the aforementioned definitions of therapist experience (confounded with training and age of therapist), school of training, internal validity variables and various sources of outcome.

TABLE 2. Variables Predictive of Effect Size

Variable	<i>F</i>	<i>p</i>	Mult. R	R ²	SR ^{2a}	Simple R	Overall <i>F</i>
1. Comparison AD ^b	3.03	.09	.55	.31	.05	.55	6.50*
2. Nonsimilar treatments used in study	2.64	.10	.57	.33	.04	.18	
3. Training in psychology (exper. therapists)	3.83	.05	.65	.42	.06	.27	
4. Solicitation of clients by experimenter (constant)	1.66	.19	.67	.44	.03	.51	
	5.30	.03					

^aSquared semi-partial correlation: unique contribution of variable to R²

^bSee text for explanation of this variable relating to experience of therapist groups.

**p* < .001, *df* = 4, 41

Table 2 shows the most useful regression beta weights (based on untransformed variables). Two variables pertaining to descriptions of comparisons between therapist groups had associated beta weights significantly different from zero in the final stepwise regression. Specifically, these categories describe comparisons between minimally-trained and untrained paraprofessionals relative to advanced, graduate-level trainees, (comparison AC), and paraprofessionals versus new and modestly experienced psychologists and psychiatrists (comparison AD). Yet alone, their respective beta weights do not make intuitive sense. Comparison AC describes a modest distance between therapist groups on experience/training and does indeed, predict negative effect sizes. This relation tends to support traditional views on the value of experience in the experience-outcome question. However, comparison AD, coded for studies having a somewhat greater discrepancy between therapist groups on experience, predicts effect size in the *opposite* (positive) direction. These results defy simple explanation. A series of partial correlations were examined to obtain a clearer look at other potentially confounding variables. It was found that the AC studies tended to utilize more technique-oriented treatments (such as behavioral). On the other hand, the group having a greater discrepancy between therapist groups on experience (AD) tended to focus on nonspecific "counseling" and less reliable, more global outcome measures.

Overall, the regression equation accounts for about 45% of the variance in effect size. The other nonsignificant beta weights in the equation were: (a) clients responding to advertisements regarding treatment, versus self-referral and nonrecruitment; (b) use of client-centered treatment, as opposed to all other treatments; (c) clients not receiving the same treatment across groups. All of these latter variables would purportedly predict *positive* rather than negative effect size, supporting the idea that less-trained therapist do as well or better than experienced clinicians. The only variable in the regression pertains to the methodological quality of studies. This variable, assignment of subjects on a rotating basis, does not make a significant and unique contribution to understanding variance in effect size.

Is there a relationship between the type of treatment and effect size? While additional categories were available for coding, only four were used often enough in studies to be included in the analyses: psychodynamic, client-centered, behavioral, and (nonspecific) counseling. Overall, no significant relationship between type of treatment and size of effect was found. If other variables are discounted, one nonsignificant trend in the data does invite speculation. Favorable outcomes for less-trained therapists relative to experienced ones are somewhat more likely to be found where "nonspecific counseling" and behavioral techniques are used, as opposed to psychodynamic therapy.

Does increasing distance between therapist groups on years of training/experience help predict effect size? The mean distance between therapist groups across these studies was 5.47 years ($sd. = 4.7$); however, there is great interstudy variability. The simple correlation between this variable and effect size is $r = -.11$ ($p < .05$). This is a modest relationship which lies in a direction predictive of better outcomes for more experienced clinicians, as distance between therapist groups being compared increases. However, the distance variable failed to account for any unique variance in effect size during examinations of the predictive power of the other variables of interest.

Does the experience-outcome relationship vary as a function of who reports outcome (i.e., the source)? It was predicted that different overall conclusions about experience and outcome would be drawn, depending on whether the patient, the therapist or an independent judge was assessing change. No such relationship was found. Taken together, the trend across studies is for a given source to report an effect size of 0. A nonsignificant tendency was for independent judges and observers to report results favoring the experienced therapist group ($p < .16$), in contrast to all other sources of outcome.

Do any client, therapist, or study variables help predict the general outcome of a study? A discriminant analysis sought to provide a look at variables that predicted whether a study produced an effect size less than zero, versus one greater than or equal to zero. In other words, what variables predict study outcomes supporting (to varying degrees) the traditional view that therapist experience is important (negative effect sizes), versus other results? The analysis entered variables into the function that passed a tolerance test for inclusion at the .001 level. The resulting discriminant function had a related eigenvalue of 4.22, a chi-square of 50.41 ($df = 17$, $p < .001$). Overall, the classification produced two distinctive groups with only two effect sizes of (46) misclassified. Group mean centroids were 1.68 for Group I (effect sizes 0 and greater) and -2.38 for Group II (negative effect sizes).

Several variables of greatest interest were tested. For example, categorizations of therapists being compared within a study (e.g. "AD," "AC": see Table 2) were used as one method of looking at the experience. No clear picture emerged for predicting the group membership of a study based on these categorizations. However, the other more continuous experience variable (involving the estimate of absolute group differences in mean years of experience) was a significant predictor of more positive outcomes for experienced therapists, relative to less-experienced ones. In addition, the discriminant function showed that outcomes tended to favor more experienced therapists when more seriously disturbed patients were seen. It must be noted however, that the range of clients was quite restricted on this dimension. Clients fell generally in the "mildly neurotic" and sometimes "character

disordered" coding categories. Patients in these studies almost never were hospitalized or had severe, chronic emotional and behavior problems.

Several other variables aided in prediction only weakly or did not make intuitive sense. These defy easy interpretation. For example, for the category of dummy variables entitled *location of treatment*, a residual group of "location unknown" (authors failing to report location) tended to be predictive of more favorable outcomes for novices.

Is there a relationship between authors' theoretical orientation and effect size? The present studies plus other available writings revealed that some authors seemed to hold a philosophy supportive of the use of nontraditional sources helping (e.g., use of paraprofessionals), of community psychology approaches to treatment (cf. Karlsruher, 1974, 1976; Cole, Oetting, & Miskimins, 1969; Poser, 1966), or emphasize the importance of nonspecific relationship factors (Strupp & Hadley, 1979). In addition, some investigators have questioned (in their other published works) whether therapist experience is of critical importance in therapy outcome (e.g., Strupp et al., 1969; Tuma, May, Yale, & Forsythe, 1978). Overall, a pattern of a failure to reject the null hypothesis regarding experience, or finding effect sizes that favor less-trained personnel is evident in the therapy outcome studies of such investigators relative to others. Numerous hypotheses accounting for this finding can be offered, such as the possible efficacy of novel versus "traditional" treatments, importance of nonspecific relationship factors (independent of training/experience), or experimenter bias.

OTHER DATA SOURCES

Some additional sources of data relating to the experience/outcome question should be considered along with the present meta-analysis. First, Stein and Holden (in press) reviewed studies pertaining to the effectiveness of minimally-trained technicians who apply behavior therapy/modification interventions. The therapists in these studies were most often college undergraduates or other persons not enrolled in formal, graduate-level training programs in psychology, psychiatry, or social work. The overall conclusion from this literature review was that the mean sizes of treatment effect for the different types of clinical problems and indices of change was comparable to that found in the large body of published literature on clinical effectiveness of behavioral strategies (Shapiro & Shapiro, 1982; Smith & Glass, 1977). Also, similar clinical populations and methodological problems were present in both bodies of studies. Since therapists in this latter literature tend to be more experienced and more intensively trained, the inference of roughly *comparable effectiveness* of minimally-trained persons relative to "professionals" can be tentatively offered. This speculation is congruent with the findings of Durlaks' (1979) review (discussed previously) and of the present meta-analysis.

Data on Dropping Out of Treatment. The findings of a recent, large scale study seems congruent with conclusion drawn in Auerbach and Johnson's (1977) aforementioned review with regard to the association between experience and the quality of the therapeutic relationship. Sue, McKinney, and Allen (1976) report on the drop-out rates of over 13,000 clients at 17 community mental health centers. They found that patients who failed to return after the initial intake session and terminated prematurely were more likely to be seen by a person having less than

graduate level training (i.e., a paraprofessional). Sue et al. suggest that the paraprofessional may cause more early termination of clients and are in need of additional training. This finding coincides with the notion that the quality of the relationship early in treatment is important and that experience does make a difference (Auerbach & Johnson, 1977).

While results are somewhat mixed, the trend in the current literature suggests that less-trained therapists tend to experience higher rates of premature dropout from treatment (cf. Fiester, 1977; Grigg, 1961; Lerner, 1972; Levitz & Stunkard, 1974; Myers & Auld, 1955; Poser, 1966; Sahaig, 1981; Slipp & Kressel, 1978). It appears that this trend is most likely evident where more traditional dynamic therapies are used in outpatient clinics (cf. Baum, Strauton, D'Zmura, & Shumaker, 1966; Dodd, 1970; Reder & Tyson, 1980).

On the other hand, studies of *college counseling centers* offering counseling services have failed to provide a similar drop out picture. Here, recent evidence suggests that level of training has no apparent bearing on whether or not clients return for counseling following the initial intake session (Betz & Shullman, 1979; Epperson, 1981; Krauskopf, Baumgardner, & Mandracchia, 1981.)

It must be understood that the reasons patients leave treatment are many. Current drop-out indices have likely included patients who improve sufficiently (after minimal contact) as well as those who no longer need treatment after some time (Fiester & Rudestam, 1975). It is not known what proportion of patients who terminate prematurely can be considered treatment failures or casualties. The concept of dropping out of treatment is a controversial one at present. Much additional research is needed before firm conclusions can be drawn regarding its use as an outcome criterion.

DISCUSSION

A relationship between experience and outcome is most likely to occur in a study when therapist groups are notably distinct on the experience dimension and where techniques other than nonspecific counseling or specific behavior techniques for circumscribed problems are the focus of study. Indeed, novices can be trained in relatively brief periods of time to provide certain circumscribed, behavioral treatments for specific problems, producing outcomes similar to those found in the literature generally. While additional research is needed, it is tentatively suggested that the level of therapist training may be related to prematurely dropping out of treatment. There seems to be an association between researchers' valuing of non-traditional, novel treatment approaches, their involvement of paraprofessionals in treatment, and favorable relative outcomes for these less-trained therapists.

In no study to date has therapist experience been shown to be of great significance in determining outcome. In some studies (e.g., Tuma et al., 1978), such variables as the use of psychotropic drugs and the therapeutic milieu were posited as more important, as have nonspecific relationship factors (Strupp & Hadley, 1979). In the present review, no variables of pronounced value were found in predicting effect size and at the same time, conceptually meaningful.

A number of other issues are highlighted by the methodology and results of available studies. Supervision is a confounding variable in current studies. Additional research is needed to assess the influence of clinical supervision, which may facilitate the therapy of novice clinicians. Therapy for the trainee should be seen

as a collaborative effort, rather than an enterprise conducted by an individual clinician. The ethical necessity of supervising novices may not have been fully considered in current authors' discussions of their research results.

In addition, other methodological problems must be grappled with in future research. The therapist's age and social status should be reasonably controlled for (as it was, for instance, in Strupp & Hadley's [1979] study). More reliable and valid operational definitions of therapist "experience" than the typically used "years of experience" or "years of training" are needed. For example, total number of cases seen or total hours of therapy contact would provide a more valid, quantitative index of experience. Improved operational definitions must consider other qualitative aspects of "experience." For instance, the difficulty of a case may bear heavily on how facilitative a given hour of therapy is in developing a clinician's level of competence.

In summary, investigators have failed too often to ask the appropriate questions, and generally have tested them with questionable designs. The results of this review should be considered to be most congruent with conclusions drawn by Auerbach and Johnson (1977) regarding the experience-outcome question: the empirical evidence supporting the assumed relationship between therapist experience and outcome is surprisingly meager.

REFERENCES

- Anthony, W. A., & Carkhuff, R. R. The functional professional therapeutic agent. In A. S. Gurman & A. M. Razin (Eds.), *Effective psychotherapy*. New York: Pergamon Press, 1977.
- Auerbach, A. H., & Johnson, M. Research on the therapists level of experience. In A. S. Gurman & M. Razin (Eds.), *Effective psychotherapy*. New York: Pergamon Press, 1977.
- Barrett-Lennard, G. T. Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs*, 1962, **76** (43, Whole No. 562).
- Baum, O. E., Stauton, B., D'Zmura, T., & Shumaker, E. Psychopathology dropouts and lower socioeconomic patients. *American Journal of Orthopsychiatry*, 1966, **36**, 629-635.
- Beck, J. C., Golden, S., & Arnold, F. An empirical investigation of psychotherapy with schizophrenic patients. *Schizophrenia Bulletin*, 1981, **7**, 241-247.
- Bergin, A. E. The evaluation of therapeutic outcomes. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change*. New York: John Wiley, 1971.
- Betz, N. E., & Shullman, S. L. Factor related to client return rate following intake. *Journal of Counseling Psychology*, 1979, **26**, 542-545.
- Bloom, L. J., & Trautt, G. M. Psychotherapists perception of clients' satisfaction following termination. *Perceptual and Motor Skill*, 1978, **46**, 1165-1166.
- Brown, R. D. Experienced and inexperienced counselors' first impressions of clients and case outcomes: Are first impressions lasting? *Journal of Counseling Psychology*, 1970, **17**, 550-558.
- Carkhuff, R. R. Differential functioning of lay and professional helpers. *Journal of Counseling Psychology*, 1968, **15**, 117-126.
- Cartwright, D. S., Kirtner, W. L., & Fiske, D. W. Method factors in changes associated with psychotherapy. *Journal of Abnormal and Social Psychology*, 1963, **66**, 164-175.
- Cartwright, R., & Vogel, J. A comparison of changes in psychoneurotic patients during matched periods of therapy and no therapy. *Journal of Consulting Psychology*, 1960, **24**, 121-127.
- Cole, E., Oetting, E., & Miskimins, R. W. Self-concept therapy for adolescent females. *Journal of Abnormal Psychology*, 1969, **74**, 642-645.
- Dodd, J. A retrospective analysis of variables related to duration of treatment in a university psychiatric clinic. *Journal of Nervous and Mental Disease*, 1970, **151**, 75-83.
- Durlak, J. A. Comparative effectiveness of paraprofessional and professional helpers. *Psychological Bulletin*, 1979, **86**, 80-92.

- Epperson, D. L. Counselor gender and early premature termination from counseling: a replication and extension. *Journal of Counseling Psychology*, 1981, **28**, 349–356.
- Feifel, H., & Eells, J. Patients and therapists assess the same psychotherapy. *Journal of Consulting and Clinical Psychology*, 1963, **27**, 310–348.
- Fiester, A. Clients' perceptions of therapists with high attrition rates. *Journal of Consulting and Clinical Psychology*, 1977, **45**, 954–955.
- Fiester, A. R., & Rudestam, K. A multivariate analysis of the early dropout process. *Journal of Consulting and Clinical Psychology*, 1975, **43**, 528–535.
- Fisher, S. G. & Nietzel, M. T. *A critique into the comparative effectiveness of professionals and paraprofessionals*. Paper presented at the Society for Psychotherapy Research Annual Convention, Pacific Grove, California, June, 1980.
- Glass, G. V., & Smith, M. L. Ask not for whom the bell tolls. *American Psychologist*, 1980, **35**, 223.
- Gomes-Schwartz, B. Effective ingredients in psychotherapy: prediction of outcome from process variables. *Journal of Consulting and Clinical Psychology*, 1978, **46**, 1023–1035.
- Grigg, A. E. Client response to counselors at different levels of experience. *Journal of Counseling Psychology*, **8**, 217–223.
- Hill, C. E. Sex of client and sex and experience level of counselor. *Journal of Counseling Psychology*, 1975, **22**, 6–11.
- Karlsruher, A. E. The nonprofessional as a psychotherapeutic agent: a review of the empirical evidence pertaining to his effectiveness. *American Journal of Community Psychology*, 1974, **2**, 61–67.
- Karlsruher, A. E. The influence of supervision and facilitative conditions on the psychotherapeutic effectiveness of nonprofessional and professional therapists. *American Journal of Community Psychology*, 1976, **4**, 145–154.
- Kirshner, L. A., Genack, A., & Hauser, S. T. Effects of gender on short-term psychotherapy. *Psychotherapy: Theory, Research and Practice*, 1978, **15**, 158–167.
- Krauskopf, C., Baumgardner, A., & Mandracchia, S. Return rate following intake revisited. *Journal of Counseling Psychology*, 1981, **28**, 519–521.
- Lerner, B. *Therapy in the ghetto: Political impotence and personal disintegration*. Baltimore: Johns Hopkins University Press, 1972.
- Levitz, L. S., & Stunkard, A. J. A therapeutic coalition for obesity: behavior modification and patient self-help. *American Journal of Psychiatry*, 1974, **131**, 423–427.
- Luborsky, L., Mintz, J., Auerbach, A., Christoph, P., Bachrach, H., Todd, T., Johnson, M., Cohen, M., & O'Brien, C. P. Predicting the outcome of psychotherapy: findings of the Penn Psychotherapy Project. *Archives of General Psychiatry*, 1980, **37**, 471–478.
- Messer, D. B., & Boals, G. F. Psychotherapy outcome in a university based psychology training clinic. *Professional Psychology*, **12**, 1981, 785–793.
- Myers, J. K., & Auld, F. Some variables related to outcome of psychotherapy. *Journey of Clinical Psychology*, 1955, **11**, 51–60.
- O'Brien, C. P., Hamm, K. G., Ray, B. A., Pierce, J. F., Luborsky, L., & Mintz, J. Group vs. individual psychotherapy with schizophrenics. *Archives of General Psychiatry*, 1972, **27**, 474–478.
- Poser, E. G. The effect of therapists' training on group therapeutic outcome. *Journal of Counseling Psychology*, 1966, **30**, 283–289.
- Reder, P., & Tyson, R. Patient dropout from individual psychotherapy: a review and discussion. *Bulletin of the Menninger Clinic*, 1980, **44**, 299–252.
- Russell, R. K., & Wise, F. Treatment of speech anxiety by cue-controlled relaxation and desensitization with professional and paraprofessional counselors. *Journal of Counseling Psychology*, 1976, **23**, 583–586.
- Sahaig, K. S. The effects of therapists' A–B Status experience and education and of client variables on psychotherapy outcome. Unpublished doctoral dissertation. California School of Professional Psychology, 1981.
- Scher, M. Verbal activity, sex, counselor experience and success in counseling. *Journal of Counseling Psychology*, 1975, **22**, 97–101.
- Shapiro, D. A., & Shapiro, D. Meta-analysis of comparative outcome studies: a replication and refinement. *Psychological Bulletin*, 1982, **92**, 581–604.
- Shelton, J. L., & Madrazo-Peterson, R. Treatment outcome and maintenance in systematic desensitization: professional versus paraprofessional effectiveness. *Journal of Counseling Psychology*, 1978, **25**, 331–335.

- Slipp, S., & Kressel, K. Difficulties in family therapy evaluation: I. A comparison of insight vs. problem-solving. II. Design, critique and recommendations. *Family Process*, 1978, **17**, 409–422.
- Smith, M. L. & Glass, G. V. Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 1977, **32**, 752–760.
- Stein, D. M., & Holden, K. Minimal training and treatment effects in behavioral interventions. *Psychological Documents*, in press.
- Strupp, H., Fox, R., & Lessler, K. *Patients view their psychotherapy*. Baltimore: Johns Hopkins Press, 1969.
- Strupp, H., & Hadley, S. Specific versus nonspecific factors in psychotherapy: a controlled study of outcome. *Archives of General Psychiatry*, 1979, **36**, 1125–1136.
- Sue, S., McKinney, H., & Allen, D. Predictors of duration of therapy for clients in the community mental health system. *Community Mental Health Journal*, 1976, **12**, 365–375.
- Sullivan, P., Miller, C., & Smelser, W. Factors in length of stay and progress in psychotherapy. *Journal of Consulting Psychology*, 1958, **22**, 1–9.
- Tuma, A. H., May, P. R., Yale, C., & Forsythe, A. B. Therapist experience, general clinical ability and treatment outcome in schizophrenia. *Journal of Consulting and Clinical Psychology*, 1978, **46**, 1120–1126.