Role of Psychosocial Treatments in Management of Schizophrenia: A Meta-Analytic Review of Controlled Outcome Studies

by Ramin Mojtabai, Robert A. Nicholson, and Bruce N. Carpenter

Abstract

This meta-analytic review sought to answer questions concerning the role of psychosocial treatments in the comprehensive management of patients with schizophrenia. The review focused on the effects of combining psychosocial treatment with somatic treatment. Findings demonstrated the additive and supplementary effects of psychosocial treatments and the durability of these effects. Patients with more chronic illness appeared to be more responsive to psychosocial treatments, as were patients in studies conducted in non-Western countries. Among the Western countries, studies from Scandinavian countries reported the least effectiveness for psychosocial treatments. There was some evidence for differential effect of psychosocial treatments on different dimensions of illness as the measures of disorganized behavior and employment showed little difference in treated and control groups. There was also some evidence for differences between different modalities of treatment as group treatments produced smaller effects. Implications for practice and future research are discussed.

Key words: Psychosocial treatment, combination treatment, psychotherapy.

***Schizophrenia Bulletin*, 24(4):569–587, 1998.**

The dramatic success of antipsychotic medication in the treatment of schizophrenia has led most clinicians to believe that medication is the treatment of choice, or at least a major component of treatment, for managing schizophrenia (e.g., Schooler and Keith 1993). However, some clinicians argue that psychosocial intervention, and more specifically psychotherapy, should be considered the treatment of choice. As recently as 1989, Karon stated that "the optimal treatment for a schizophrenic is psychotherapy, from a competent therapist, without medication, if the patient, the therapist, and the setting can tolerate it" (p. 146).

Sigmund Freud, the forefather of modern psychotherapy, expressed pessimism about psychoanalytic treatment of schizophrenia (Fenichel 1945). However, in the years that followed, many therapists attempted psychological treatment for these patients (Fromm-Reichmann 1950; Arieti 1955). These treatments were based on psychogenic theories of etiology and sought to cure the illness by addressing its generative mechanisms. The advent of antipsychotic medications and developments in the biological understanding of schizophrenia on the one hand, and lack of supporting evidence for some of the psychogenic theories such as "schizophrenogenic mother" on the other, led to pessimism concerning the effectiveness of the psychological treatments for schizophrenia (Bellack and Mueser 1993).

However, better understanding of the limitations of antipsychotic medications and the emergence of interactive etiological models of schizophrenia (e.g., stress-diathesis model) in more recent years have contributed to a renewal of interest in psychological interventions (Schwartz et al. 1993). Beyond the old controversies of psychotherapy versus medication, a new perspective seems to be emerging that psychosocial interventions can be used beneficially in conjunction with medication and that the combination may actually have an additive or synergistic effect (e.g., Carpenter and Keith 1986; Schwartz et al. 1993). The emergence of more focused approaches to psychosocial treatment (such as social skill training and cognitive training), which attempt to remedy deficits in a specific area of functioning, has also contributed to renewed interest in psychosocial interventions in general.

One of the characteristics that distinguishes many of the more recent approaches to psychosocial treatment is that they set more modest goals. Where earlier psy-

Reprint requests should be sent to Dr. R. Mojtabai, Division of Epidemiology and Community Psychiatry, Department of Psychiatry, Columbia University, 601 W. 168th St., Suite #32, New York, NY 10032.

chotherapists espoused a curative role for psychotherapy (Arieti 1955), many of the more recent psychosocial therapists limit their attempts to removal of the ill effects of the disorder and prevention of relapse. Viewing the combination of medication and psychological treatment as optimal, advocates of the more recent psychosocial approaches aim to address aspects of the patients' lives that pharmacological treatment cannot.

The beneficial effects of combining psychosocial treatments and medication are supported by the results of controlled studies (Schooler and Keith 1993) and by the clinical experience of many clinicians (Sarti and Cournos 1990). However, the exact nature and magnitude of these effects have not been adequately investigated. For example, Falloon and Liberman (1983) have suggested a synergistic effect for those psychosocial treatments, such as family interventions, that enhance the coping capacity of the patient's support system. Other reviewers have suggested an additive effect for psychosocial treatments in general (Schooler and Keith 1993).

Another unresolved issue is the comparative efficacy of different psychosocial treatments. A number of recent reviewers report that certain forms of psychosocial treatment are more effective than others. Family interventions and social skills training programs are among the favored treatments, while individual psychodynamic treatments are viewed less favorably (Bellack and Mueser 1993; Schooler and Keith 1993; Schwartz et al. 1993). Consistent with this belief, Mueser and Berenbaum (1990) proposed a moratorium on the use of psychodynamic treatments for schizophrenia.

Yet another important and unresolved issue is the impact of the characteristics of the patients (e.g., gender, chronicity of illness) and of the treatments (e.g., duration, frequency) on the results of psychosocial treatment (Bellack and Mueser 1993). Finally, the question of durability of effects needs to be addressed. Bellack and Mueser (1993) argue that no one expects neuroleptic medications to continue to be effective after they have been discontinued, so expecting psychosocial treatments to be effective after termination is not justified. According to these authors, assuming durability as a valid test for psychosocial interventions is inconsistent with the chronic nature of schizophrenia. However, some forms of psychosocial treatment are claimed to make basic and durable changes in patients' functioning. Followup studies constitute an important test for the construct validity of such treatments. Furthermore, evidence of durability affects decisions on continuation of treatment and on the need for booster sessions. In any case, the fact that a considerable number of investigators attempt followup investigations is evidence of continued interest in this issue.

Answering these and many other questions concerning the effects of psychosocial treatments for schizophrenia has important implications in terms of both treatment planning for individual patients and for large-scale mental health policies. Future researchers need to address these issues. However, previous studies contain a wealth of information that can be used to inform and guide future trials. Several excellent narrative reviews of this literature have recently been published (Bellack and Mueser 1993; Schooler and Keith 1993; Schwartz et al. 1993). But, the number of published primary studies is so large that narrative reviews can include only a small portion of such studies. Quantitative review approaches, generally known as meta-analyses, are more capable of handling larger numbers of primary studies, and by extracting standardized measures of effectiveness (effect size), these approaches enable the reviewers to compare the effects of different classes of treatments, patients, or outcome measures more objectively. In addition, by systematic assessment of factors that vary across individual studies, meta-analysis is capable of answering questions that cannot be addressed easily in individual studies (e.g., the effect of facility or geographical location in which the treatment is administered).

Four previous meta-analyses have examined the efficacy of psychosocial treatments for schizophrenia. However, none of these studies was a general and comprehensive review of the relevant literature. The meta-analyses by Smith et al. (1980) reported an added effect for psychosocial treatments when used in conjunction with medications in patients with schizophrenia as well as patients with other diagnoses. However, these authors included only a relatively small sample of primary studies for each diagnostic group. In addition, the Smith et al. study suffered from methodological problems, such as the use of individual outcome measures as units of analysis (Robinson et al. 1990).

Since Smith et al. (1980), three quantitative reviews of psychosocial treatments for schizophrenia have been published (Quality Assurance Project 1984; Benton and Schroeder 1990; de Jesus Mari and Streiner 1994). The Quality Assurance Project (1984), which reviewed 26 primary studies, reported that a group of treatments labeled as "social interventions" had a synergistic effect in combination with medications, but another group of treatments labeled simply as "psychotherapy" did not have any effect, either alone or in combination with medications. The social interventions were characterized by an emphasis on family education and development of a social network. However, it is not clear from the review how the psychotherapy group was characterized. In addition, this review suffered from the same methodological problems as the Smith et al. (1980) study.

The review by Benton and Schroeder (1990) examined the effect of social skills training in schizophrenia. These authors included 27 studies in their review and reported a posttreatment effect size of 0.65 on average. They also attempted to examine the impact of different moderator variables on the effectiveness of treatment but found no significant results, perhaps due to the small number of studies. This review was free from the methodological problems of Smith et al. (1980) and the Quality Assurance Project (1984). However, Benton and Schroeder aggregated two types of studies: one type in which social skills training was compared with another form of psychosocial treatment and the other type in which the comparison was with an empty control or standard treatment. These two types of studies measure different parameters. Also, the studies included were not limited to patients with a diagnosis of schizophrenia only. (However, Benton and Schroeder did not find a difference in effect size between studies that included patients with schizophrenia only and those that included other patients as well.)

Most recently, de Jesus Mari and Streiner (1994) reviewed the effect of family interventions in schizophrenia. They reported a significant reduction in relapse rate in the experimental group. However, their review was limited to only six primary studies and the outcome measure of interest was relapse rate. These reviewers admitted that, as a primary measure of outcome, relapse rate has many limitations.

Only limited conclusions can be drawn from the previous reviews as to the role and significance of psychosocial treatments for schizophrenia and the impact of different factors on the results of these interventions. The purpose of the present review is to provide a comprehensive quantitative summary of the controlled outcome studies. Specifically, an attempt was made to answer the following questions with the use of a large pool of published studies and careful application of meta-analytic method: (1) Does the addition of a psychosocial treatment to a standard medical regimen enhance treatment outcome? And if so, what is the magnitude of this added effect? (2) How do different moderators (in particular, modality and orientation) affect the results of psychosocial treatment? (3) How durable are the results of psychosocial treatment?

Methods

Studies. Inclusion was limited to studies that reported on patients with a diagnosis of schizophrenia only or in which the effects of treatment could be estimated separately for these patients. To be included, the study had to report results from a comparison of two or more groups of patients, at least one of which received a form of psychosocial treatment. Studies in which two forms of psychosocial treatment were compared were included, as were studies in which psychosocial treatment was compared with a form of somatic treatment for schizophrenia, including antipsychotic medications and electroconvulsive therapy (ECT). Studies in which one form of psychosocial treatment was used in conjunction with a form of somatic treatment was compared with another form of treatment or another combination of treatments were also included.

The definition of psychosocial treatment used for this review was broad. It included traditional psychotherapy as well as newer forms of psychological treatment (e.g., family psychoeducation, cognitive training) and social intervention such as community treatment. Forms of psychological treatment that are of little direct clinical interest (e.g., training patients on the Wisconsin Card Sorting Test and then testing their performance on the same test [Heaton 1981]) were not included. As Bellack and Mueser (1993) note, these studies are conducted to demonstrate the plasticity of test performance and not to produce a clinical change. The only exception to this criterion were studies with a clinically relevant psychological treatment used in conjunction with such experimental interventions or a clinically relevant outcome measure used in the study of a clinically nonrelevant intervention.

Using these criteria, two computerized databases were searched for relevant studies: PSYCHLIT for years 1974–94 and MEDLINE covering years 1966–94. Additionally, recent issues of the journals that had published most of the identified studies were manually searched.[1] A fourth source of studies was the reference section of previous reviews. Overall, 200 papers, books, and book chapters met the inclusion criteria and were obtained. Of these sources, 59 did not meet the criteria for inclusion in the study: 17 sources reported preliminary or less detailed findings that were later reported as final results or in more detail in other available sources; 13 sources reported on interventions of little clinical interest (see above); 10 sources reported results from studies with no comparison group; 9 sources either did not specify the diagnosis of patients and referred to them as "psychotic" patients or indicated that patients with other diagnoses besides schizophrenia were included; 7 sources did not include enough data for calculation of effect size; and 3 sources reported on the results of process measures and

[1]The manual search was conducted for the following journals: *Archives of General Psychiatry, American Journal of Psychiatry, British Journal of Psychiatry, Hospital and Community Psychiatry, Schizophrenia Bulletin, Psychopharmacology Bulletin,* and *Acta Psychiatrica Scandinavica.*

included no outcome measures. After excluding these sources, 141 sources reporting on 106 individual studies remained and were included in the study.[2]

In the selection of studies, no *a priori* criteria of quality such as internal validity were applied. Whether such characteristics affect the results is an empirical question that can be answered objectively in a meta-analytic review of a body of literature (Smith et al. 1980). The studies were coded for such characteristics as random assignment, use of treatment manual, and blindedness of outcome measures, and the impact of these variables on the results of the studies were examined.

The studies were also coded for the different characteristics of the patients and treatments. Descriptive characteristics of the 106 studies included in the review are presented in table 1. Most studies were from inpatient settings (n = 65, 61.3%); 20 (18.9%) were from outpatient settings, and 4 (3.8%) from partial hospitalization settings. In 2 (1.9%) studies, treatment began in an inpatient setting and continued on an outpatient basis; for 15 (14.1%) studies the setting was not specified. Studies reviewed were from different parts of the world: 65 (61.3%) studies were from the United States and Canada, 10 (9.4%) from Great Britain, 8 (7.5%) from the Scandinavian countries, 14 (13.2%) from Continental Europe, and 9 (8.5%) from non-Western countries, including China and Israel.

Studies also differed in terms of the facilities in which they were conducted: 25 (23.6%) were from university hospitals or research centers, 13 (12.3%) from Veterans Affairs (VA) hospitals, 17 (16.0%) from other types of public hospitals (e.g., State, municipal), and 1 (0.9%) from a private institution. For the remainder of the studies, either the facility was defined as a medical center (n = 7, 6.6%) or its nature was not specified (n = 43, 40.6%).

The studies in this review spanned a rather long time period, as well as a wide range of geographical locations and settings. Therefore, it is not surprising that they used different definitions of schizophrenia and different approaches to diagnostic decision making (e.g., structured interview, criteria, clinician's judgment). Unfortunately, a large number of studies did not report specific criteria (n = 53, 50%). Of the studies that did provide this information, nine (8.5%) used *DSM–III* criteria (American Psychiatric Association 1980), nine (8.5%) used *DSM–III–R* (American Psychiatric Association 1987), eight (7.5%) used Research Diagnostic Criteria (RDC; Spitzer et al. 1978), seven (6.6%) employed International Classification of Diseases, 8th edition (ICD–8; World Health Organization 1967) or 9th edition (ICD–9; World Health Organization 1978), seven (6.6%) used the Present State Exam (PSE; Wing et al. 1974), two (1.9%) listed *DSM–II* (American Psychiatric Association 1968), and one (1%) used Feighner's criteria (Feighner et al. 1972). The other 10 studies (9.4%) used other less well-known criteria or reported a set of criteria for the purpose of their study.

Only 46 studies (43.4%) specified the diagnostic approach they employed. The diagnostic decision was based on clinical judgment of the examiner from an unstructured interview in 25 (23.6%) studies; 9 of these (8.5%) reported the use of a set of criteria, and 16 (15.1%) did not. In 11 (10.4%) studies, diagnostic decision was based on structured interview, and in 10 (9.4%) it was based on a review of the patient's chart.

Estimating Treatment Effects. Outcomes reported in the studies reviewed were translated into Cohen's (1977) d, a standardized estimate of effect size. Cohen's d is defined as

$$d = \frac{m_1 - m_2}{s}$$

Where m_1 and m_2 are the means of the treatment and control groups, respectively, and s is the pooled within-group standard deviation (SD). Thus, Cohen's d expresses the difference between means relative to within-group variation. For example, a d of 1.0 indicates that the mean of the treatment group is one SD higher than the mean of the control group. As Glass et al. (1981) note, associating regions of the effect size metric with descriptors such as "small" and "moderate" is not justifiable since "dissociated from the context of decision and comparative value, there is no inherent value to an effect size of 3.5 or 0.2" (p. 104). However, a few examples of effect size estimates from other treatments may put the estimates obtained in the present study in perspective. In their comprehensive review of 156 meta-analyses of psychological, educational, and behavioral treatments, Lipsey and Wilson (1993) estimated an average d of 0.47. In comparison, different medical or surgical treatments for physical illness (e.g., coronary bypass surgery, treatment of arthritis, etc.) produce effect sizes in the range of 0.08 to 0.47, and physical treatments for psychological disorders (e.g., electroconvulsive therapy [ECT] for depression, drug treatment for hyperactivity, etc.) produced effect sizes in the range of 0.11 to 0.96 (Lipsey and Wilson 1993).

Hedges has demonstrated that d calculated for small samples is biased. In our estimation of d, we corrected this bias according to Hedges' correction formula (Hedges and Olkin 1985, p. 81, equation 10). A large number of the effect sizes were calculated from means and SDs

[2]The list of references for the studies included in the meta-analysis is available from the first author.

Table 1. Characteristics of studies and patients

Characteristics	Number of studies	Mean	Range
Characteristics of the studies			
Publication year	106	1979.4	1954–1994
Number of patients	106	63.4	10–374
Demographic characteristics of patients			
Percent female	87	34.7	0–100
Mean age	93	34.6	16.0–66.0
Percent married	32	23.1	0.0–56.0
Mean years of education	24	10.8	8.2–13.1
Psychiatric history of patients			
Mean years since onset of illness	46	8.5	0.16–23
Mean age at onset	43	25.3	19.1–46.8
Mean number of previous hospitalizations	45	2.4	0.0–4.9
Psychiatric-psychological characteristics of patients			
Percent with drug/alcohol comorbidity[1]	24	0.9	0.0–21.0
Percent with paranoid subtype	17	34.2	10–73
Mean IQ score	8	96.8	92.7–101
Mean BPRS baseline score[2]	13	43.4	31–52

[1]Drug or alcohol abuse was one of the exclusion criteria for recruiting patients in many studies.
[2]BPRS = Brief Psychiatric Rating Scale (Overall and Gorham 1962).

reported in the original studies. When this information was not available, we used other statistics to estimate the effect size (Glass et al. 1981, chapter 5; Nicholson and Berman 1983). For dichotomous outcome measures (e.g., rehospitalization), probit transformation was performed. When primary researchers reported a result only as failing to reject the null hypothesis, we followed the common practice of estimating the effect size conservatively as zero (e.g., Shadish et al. 1993). In a few cases, the researchers described an outcome measure but did not report results from it or reported results as nonsignificant. Excluding these measures would have inflated the effect size estimates because investigators are more likely to report results for measures that produced reliable group differences. In such cases we followed the conservative approach of estimating the effect size as zero (Robinson et al. 1990). As the analyses in the following sections demonstrate, this practice of estimating nonsignificant and nonreported results did not change the estimates of effect size considerably.

In general, mean effect sizes were calculated for each study by averaging the effect sizes of all outcome measures from that study. However, different outcome measures used in a study have different sources and reflect different constructs. Therefore, we also calculated mean effect sizes separately for different sources of outcome measures (e.g., self-rated measure, other-rated measure, etc.) and for different contents of outcome measures (e.g., positive symptoms, cognitive functioning, etc.). As the results of regression analysis (see below) demonstrate, content of outcome measure affected the effect size estimates whereas the source of outcome measure did not.

Mean effect sizes for classes of studies and confidence intervals (CIs) for these estimates were calculated by the method described by Hedges (1994, pp. 286–289) in which each effect size is weighted by the inverse of its conditional variance, a function of the sample size. Between-groups heterogeneity statistics (Q_B) were used in comparisons across classes of studies. Within-group heterogeneity statistics (Q_W) were used to determine which classes were the major sources of within-group heterogeneity and which groups had relatively homogeneous effect sizes (Hedges 1994, pp. 289–290, equations 19-10 and 19-11). Heterogeneity statistics have a similar interpretation to variance in analysis of variance. The Q_B is the weighted sum of squares of group means about the grand mean, and the Q_W is the total of the weighted sum of squares of the individual effect estimates about the respective group means. The sum of Q_B and Q_W is the total heterogeneity statistic (Q_T), which is the weighted total sum of squares about the grand mean.

Preliminary Analyses. While the majority of the studies included only one comparison between two groups, 34 (32.1%) studies included more than one comparison. In all, the 106 studies included 172 comparisons (mean = 1.62, range = 1–7). These comparisons could be classified into seven distinct types: (1) psychosocial treatment plus somatic (or standard) treatment compared with somatic or standard treatment alone (103 comparisons

from 71 studies); (2) combined treatment compared with no treatment (10 comparisons from 5 studies); (3) combined treatment compared with psychosocial treatment alone (6 comparisons from 5 studies); (4) psychosocial treatment only compared with no treatment (9 comparisons from 6 studies); (5) somatic treatment compared with no treatment (12 comparisons from 6 studies; this type of comparison emerged only in studies that also included other comparisons); (6) psychosocial treatment compared with somatic treatment (4 comparisons from 3 studies); and (7) two different forms of psychosocial treatment, each in combination with somatic treatment, compared with each other (28 comparisons from 23 studies). As may be noted, the total number of studies is higher than 106 because some studies included multiple comparisons of different types.

We aggregated multiple effect sizes within comparisons to the comparison level by adapting a method reported by Robinson et al. (1990) and Shadish et al. (1993). Because comparisons within a study always used a common control group, results from such comparisons cannot be considered independent. Therefore results from these comparisons were aggregated and an average effect size was calculated for each study. A between-group test suggested that different types of comparisons estimated different effect sizes (Q_B = 252.20, degrees of freedom [*df*] = 5, $p < 0.001$) (table 2). For this analysis we excluded type 7 for which the sign of effect size estimate is arbitrary, depending on which treatment is considered the experimental treatment and which one, the control. As noted before, we also aggregated effect sizes from similar types of comparisons in the same studies. As a result of these manipulations, only 96 of the original 172 comparisons could be included in analyses reported in table 2.

Studies comparing combined treatment with no treatment (type 2) produced effect sizes larger than studies comparing each one of these treatments to no treatment (types 4 and 5). The evidence for added benefit of psychosocial treatment in these studies is especially important because in all three types, the treatments were compared to the same form of control group (i.e., no treatment or empty control).

As the results of this preliminary analysis demonstrate, effect sizes from different types of studies are heterogeneous and cannot be aggregated. Therefore, we limited our analyses exclusively to type 1, which includes the largest number of studies. Results from these studies have the most ecological validity and most closely represent the "real world," where somatic treatments and especially antipsychotic medications have become a mainstay of treatment. The large number of type 1 studies is also evidence of their relevance. In analyzing this type of study, the question of interest is whether psychosocial treatment adds to the benefits of somatic (or standard) treatment. Standard treatment, for the comparisons in this study, included various forms of inpatient and outpatient treatment. In most cases of such comparisons, the authors mentioned the use of somatic treatments as standard treatment, but in a few cases they specified psychosocial treatments and referred to them as standard treatment (milieu therapy in five comparisons and supportive psychotherapy in two comparisons). Wherever the authors referred only to "standard" or "routine" treatment, it was assumed that a form of somatic treatment had also been used, unless otherwise noted.

The somatic treatment in most cases was antipsychotic medication; in three studies it was ECT. The mean dosage (in chlorpromazine equivalents) could be calculated for the experimental group in 10 studies (497.8 mg,

Table 2. Effect size estimates from different types of comparisons

Type of comparison	d_+	95% CI	Q_W[1]	*p*
Psychosocial treatment + somatic (or standard) treatment vs. somatic (or standard) treatment (71)	0.39	$0.32 < \delta < 0.44$	172.28	<0.001
Psychosocial treatment + somatic (or standard) treatment vs. no treatment (5)	0.85	$0.62 < \delta < 1.09$	12.28	0.015
Psychosocial treatment + somatic (or standard) treatment vs. psychosocial treatment alone (5)	0.27	$0.03 < \delta < 0.51$	8.53	0.074
Psychosocial treatment alone vs. no treatment (6)	0.23	$0.02 < \delta < 0.45$	5.01	0.415
Somatic treatment alone vs. no treatment (6)[2]	0.37	$0.19 < \delta < 0.55$	5.01	0.415
Psychosocial treatment alone vs. somatic treatment alone (3)	–0.06	$-0.32 < \delta < 0.21$	12.93	0.002

Note.—Numbers in parentheses indicate the number of studies that included the specified type of comparison. Studies comparing two forms of psychosocial treatment are not reported in this table since in this category the sign of effect size estimate is arbitrary, depending on which treatment is considered the experimental treatment and which one, the control. In addition, comparisons of the same type within each study were aggregated. Therefore, the number of comparisons in this table is only 96. CI = confidence interval.

[1]Q_w is a measure of within-study heterogeneity. Statistically significant values indicate that the effect sizes from studies within a class are heterogeneous.

[2]This type of comparison emerged only in studies that also included other comparisons.

SD = 307.3 mg). Nine of these studies reported a mean dosage for the control group (514.4 mg, SD = 316.6) indicating no difference between the two groups (paired t = 0.28, $df = 8$, $p = 0.78$).

Reliability of Codings. To verify the reliability of the coding scheme, two independent coders recoded 14 randomly selected studies (13%) according to the instructions from a coding manual. Interrater reliabilities for these codings were calculated with kappa for categorical variables and intraclass correlation coefficient (ICC) for continuous variables (Bartko and Carpenter 1976; Maclennan 1993). For a number of variables the base rate of certain ratings was unusually high. In such cases, disagreement on only a few cases would reduce kappa considerably. The resultant kappa should therefore be regarded with extreme caution.

Interrater reliability estimates for most study characteristics (e.g., publication date, number of subjects, country, basis for diagnostic decision, diagnostic criteria used, setting, type of facility, randomization, manualization, etc.) were high (0.81–1.0) (Landis and Koch 1977). Kappa for rating of type of comparison was 0.82, and ICCs for the two patient variables of age and male to female ratio were perfect.

Kappa calculated for the variable of authors' allegiance was only 0.44, in the moderate range of agreement. However, of 14 pairs of ratings for this variable, only 2 pairs showed disagreement. In 11 cases, both raters agreed that the authors had an allegiance for the experimental treatment, and in 1 case both agreed that the authors did not have any allegiance. Kappa for the ratings of the therapists' allegiance was only 0.22, in the fair range of agreement. For this variable, 6 of the 14 pairs of ratings disagreed.

Some characteristics of the treatments, such as the number of sessions, frequency of sessions, duration of treatment, theoretical orientation of treatment, etc., proved to have high reliability (0.81–1.0); whereas, modality of treatment, content of outcome measures, and source of outcome measures had kappas of 0.71, 0.67, and 0.65, respectively, all in the range of substantial reliability. Effect size calculation had an ICC of 0.99.

Results

General Descriptive Results. Effect sizes for 71 studies comparing combination treatment to somatic (or standard) treatment alone ranged from –0.48 to 2.23 and were positively skewed (figure 1). The effect size of 2.23, derived from the study by Beal et al. (1977), was an outlier and was excluded from further analyses, reducing the sample size to 70 studies. The weighted least squares (WLS) average of the effect sizes was $d_+ = 0.39$ with a standard error of 0.030 and a 95 percent CI of $0.32 < \delta < 0.44$. We conservatively set to zero the effect sizes from all measures that reported the results as nonsignificant and

Figure 1. Stem and leaf display of 71 effect sizes from studies comparing a form of psychosocial treatment combined with somatic (or standard) treatment to somatic (or standard) treatment alone

Stem	Leaf
2.2	3
2.1	
2.0	
1.9	
1.8	
1.7	
1.6	
1.7	
1.6	
1.5	2, 3, 7
1.4	
1.3	
1.2	2
1.1	0
1.0	1, 8
0.9	1, 5, 7, 9
0.8	2
0.7	2, 3, 5, 6, 8, 9
0.6	1, 2, 4, 7, 8
0.5	1, 1, 4, 6, 6, 6, 8, 8, 9, 9
0.4	0, 2, 3, 3, 4, 7, 7
0.3	0, 2, 4, 6, 7
0.2	0, 1, 1, 1, 3, 3, 4, 7, 8
0.1	0, 4, 6, 8
0.0	0, 0, 4, 4, 5, 7, 8
-0.1	4
-0.2	2, 7
-0.3	5
-0.4	8
-0.5	

The integer before decimal and the first decimal place in the effect size calculated from each study is represented as the "stem" and the second decimal place as a "leaf." The number of leaf entries represents the number of studies within a certain range. Thus, the topmost entry represents the one study that produced an effect size in the 2.2 to 2.3 range ($d = 2.23$); three studies produced effect sizes in the 1.5 to 1.6 range ($d = 1.52$, 1.53, and 1.57), etc.

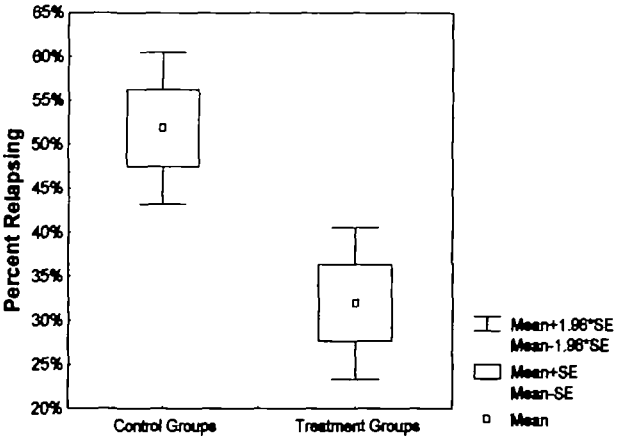
those that were used in the study but for which no results were reported. To assess the impact of this approach on the resultant estimate of average effect size, we calculated another estimate excluding such outcome measures. Of the 917 outcome measures reported for these studies, the effect sizes estimated for 73 (8%) were set to zero for the above reasons. The WLS average effect size calculated after exclusion of these measures was less than 10 percent larger ($d_+ = 0.42$) and within 95 percent CIs for the original estimate. This finding is consistent with results from other meta-analytic studies (e.g., Shadish et al. 1993). We decided to use the more conservative approach of including these zero effect sizes in all the analyses. However, we examined the impact of this decision on our substantive conclusions by entering the number of such measures as a variable in the regression analyses below.

An effect size of 0.39 implies that a typical patient in the experimental treatment group was better off than 65 percent of control patients. However, a test of heterogeneity of effect sizes was statistically significant, ruling out homogeneity of effect size estimates from different studies in this broad category ($Q_W = 168.66$, $df = 69$, $p < 0.001$). Therefore, we further classified the studies according to different categorical moderator variables and calculated effect sizes and heterogeneity statistics for these classes. Results of these analyses are reported in the section on moderator variables.

Effect of Psychosocial Treatment on Relapse. To highlight the clinical relevance and significance of the results, we also analyzed the effects of psychosocial treatment on relapse, which is of practical importance in the management of schizophrenia. We calculated Cohen's *d* through probit transformation of relapse frequencies and included this measure in our calculation of mean effect size for each study. We also examined this effect using raw relapse frequencies. Fourteen studies reported the effects of treatment on relapse. As in previous analyses, frequencies from comparisons in the same studies were aggregated and mean relapse frequencies for experimental and control groups were derived for each study. Although rehospitalization was the major outcome measure for relapse, three studies reported symptomatic relapse as measured by an other-rated instrument. The majority of studies reported on relapse during the outpatient treatment. Two studies reported on relapse during followup, and one study spanned both treatment and followup. The median period of time during which relapse was recorded was 17 months. Figure 2 presents relapse frequencies for treatment and control groups from the 14 studies.

Relapse frequencies for patients who received psychosocial treatment in addition to somatic (or standard) treatment were consistently lower than for patients who received only somatic (or standard) treatment. The relapse frequencies for the psychosocial treatment groups were, on average and after weighting for sample size, 20 percent lower than that for the control groups.

Figure 2. Relapse frequencies in the treatment and control groups for 14 studies



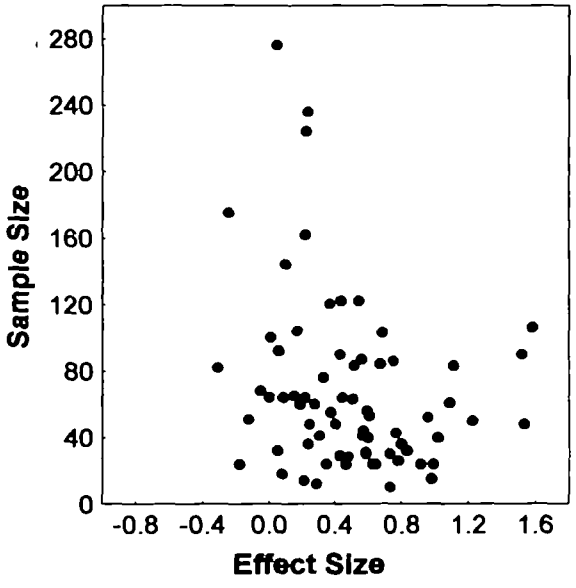
SE = standard error.

Moderator Variables. We examined the relationship between moderator variables and the outcome of the studies by calculating WLS correlations for continuous variables and between-group heterogeneity statistics for the categorical variables. In all, the effects of 35 different study, patient, and outcome variables were examined. This needs to be kept in mind when interpreting the results of statistical significance testing. Some authors have suggested a method of adjusting the alpha level for multiple testing (Grove and Andreasen 1982). The adjusted alpha level for *n* tests might be calculated by dividing the originally adopted alpha level by *n*. For the present study, we originally adopted an alpha level of 0.05, therefore the adjusted alpha level would be 0.05/35 or 0.00143. However, other authors (e.g., Bartko et al. 1988) have warned that this adjustment may be too stringent. In any case, we have reported the actual *p* values larger that 0.001 and have left the decision whether to use adjustment to the reader. It is reassuring that the use of adjusted alpha levels had little impact on most of the substantive conclusions from this review.

Among the different study variables, publication date and sample size each had a statistically significant impact, whereas publication form and the degree of the first author had no appreciable impact. More recent studies tended to produce larger effect sizes ($r = 0.29$, $df = 68$, $p = 0.015$) before adjustment for multiple testing. However, this effect was not statistically significant after adjust-

ment. Studies with larger sample sizes produced smaller effect sizes ($r = -0.38$, $df = 68$, $p = 0.0013$), an association statistically significant before and after adjustment for multiple testing. This finding may reflect a publication bias (Begg 1994) insofar as studies with small sample sizes that do not produce statistically significant results are less likely to get published (figure 3).

Figure 3. Relationship of sample size and effect size for 70 studies



The quality of the studies was operationally defined by such variables as random assignment, manualization, equal attrition rates, use of structured interview in diagnosis, and control for patient expectation. The effect of patient expectation could not be examined because none of the studies reported this variable. None of the other variables had any reliable effects on the outcome of the comparisons.

Allegiance of the authors was coded based on a set of objective criteria, including the explicit endorsement of a treatment approach and prior publications by the authors. Effect sizes from those studies in which authors had a clear allegiance to the experimental treatment were larger than those in which the allegiance was not clear ($Q_B = 7.21$, $df = 1$, $p = 0.007$; table 3). This effect was statistically significant before adjustment for multiple testing, but not significant after adjustment.

The impact of the source and content of outcome measures has been documented in previous meta-analyses (e.g., Smith et al. 1980). We coded individual effect sizes on the basis of source and content and calculated separate effect sizes for these classes. Three classes of outcome measures were coded according to source (self-rated vs. other-rated vs. objective measures [e.g., total days spent in hospital]) and 11 classes according to content (e.g., negative symptoms, behavioral disorganization; table 4). Direct comparison of effect sizes in these classes via between-groups tests of heterogeneity would not produce reliable results as a number of the measures in different

Table 3. Effect size estimates as a function of different predictor variables

Variable	d_+	95% CI	Q_W[1]	p
Facility				
University hospital/research center (14)	0.48	$0.33 < \delta < 0.63$	37.18	<0.001
Public hospital (non-VA) (10)	0.36	$0.22 < \delta < 0.50$	12.63	0.180
VA hospital (12)	0.22	$0.09 < \delta < 0.35$	19.78	0.048
Medical center (5)	0.51	$0.31 < \delta < 0.70$	23.29	<0.001
Other/not specified (29)	0.44	$0.33 < \delta < 0.54$	63.66	<0.001
Countries				
USA or Canada (38)	0.29	$0.22 < \delta < 0.36$	54.52	0.031
Great Britain (6)	0.55	$0.28 < \delta < 0.81$	2.02	0.846
Continental Europe (12)	0.57	$0.40 < \delta < 0.73$	29.39	0.002
Scandinavian countries (6)	0.05	$-0.15 < \delta < 0.25$	13.04	0.016
Other countries (non-Western) (8)	0.92	$0.73 < \delta < 1.10$	18.29	0.011
Diagnosis				
Based on objective criteria—narrow (14)	0.61	$0.45 < \delta < 0.77$	30.22	0.004
Based on objective criteria—broad (15)	0.54	$0.40 < \delta < 0.67$	30.31	0.007
Not based on objective criteria (41)	0.30	$0.23 < \delta < 0.37$	89.98	<0.001
Authors' allegiance for experimental treatment				
Authors have an allegiance (50)	0.44	$0.37 < \delta < 0.52$	129.36	<0.001
Author's allegiance not clear (20)	0.28	$0.18 < \delta < 0.38$	32.09	0.031

Note.—Numbers in parentheses indicate the number of studies. CI = confidence interval; VA = Veterans Affairs.

[1] Q_W is a measure of within-study heterogeneity. Statistically significant values indicate that the effect sizes from studies within a class are heterogeneous.

Table 4. Effect size estimates for outcome measures from different sources and with different contents

Source/outcome	d_+	95% CI	Q_W[1]	*p*
Source of outcome measures				
Self-rated (11)	0.21	$0.04 < \delta < 0.37$	9.91	0.449
Other-rated (60)	0.39	$0.33 < \delta < 0.46$	165.84	<0.001
Objective measures (35)	0.38	$0.30 < \delta < 0.46$	130.73	<0.001
Content of outcome measure				
Symptomatology				
Positive symptoms (9)	0.32	$0.11 < \delta < 0.53$	5.99	0.648
Negative symptoms (24)	0.51	$0.41 < \delta < 0.62$	132.18	<0.001
Thought disorder (12)	0.59	$0.40 < \delta < 0.78$	51.14	<0.001
Anxiety/depression (15)	0.26	$0.12 < \delta < 0.39$	16.92	0.261
Disorganized behavior (19)	0.13	$0.00 < \delta < 0.25$	46.45	<0.001
General symptomatology (12)	0.26	$0.08 < \delta < 0.45$	5.38	0.911
Cognitive functioning (12)	0.36	$0.18 < \delta < 0.54$	18.35	0.070
Objective measures of adjustment (e.g., days working) (20)	0.26	$0.15 < \delta < 0.37$	95.59	<0.001
Relapse (including rehospitalization, symptomatic relapse, total days spent in hospital after index episode, time to rehospitalization) (21)	0.46	$0.35 < \delta < 0.57$	48.87	<0.001
Employment (8)	0.22	$0.04 < \delta < 0.39$	47.68	<0.001
Compliance with medication (4)	0.50	$0.16 < \delta < 0.84$	1.03	0.794

Note.—Numbers in parentheses indicate the number of studies that used outcome measures with the specified source or content. CI = confidence interval.

[1] Q_w is a measure of within-study heterogeneity. Statistically significant values indicate that the effect sizes from studies within a class are heterogeneous.

classes come from the same studies and therefore cannot be regarded as independent. However, we were able to test their impact via a multiple regression analysis, as reported in a following section.

Among the different patient variables, only the variable of time since onset of illness had a statistically significant effect before and after adjustment for multiple testing ($r = 0.63$, $df = 30$, $p < 0.001$). The more chronic the illness was, the larger the effect size tended to be. Other factors such as patients' gender, age, marital status, education, IQ score, alcohol/drug abuse, and even previous hospitalization did not have any reliable effect.

We classified studies into those that had used a formal set of diagnostic criteria and those that had not. On comparison, studies that had used formal criteria in diagnosis of patients produced larger effect sizes ($Q_B = 18.15$, $df = 1$, $p < 0.001$; table 3). This effect was statistically significant before and after adjustment for multiple testing. In addition, we classified the diagnostic criteria as reflecting either a narrow or a broad definition of schizophrenia (see Hegarty et al. 1994) based mainly on whether the criteria required a minimum of 6 months' duration. In a previous meta-analysis, criteria that include such a requirement (narrow definition) tended to include a group of patients with a less favorable prognosis (Hegarty et al. 1994). This distinction had no reliable effect on the results of studies in our sample ($Q_B = 0.49$, $df = 1$, $p = 0.484$). However, it should be noted that most study samples were composed of chronic patients (as evidenced by the demographics of the patients), and for such patients the minimum of 6 months' duration may be less relevant. We also assessed whether the distinction between paranoid and nonparanoid subtypes had an impact on treatment outcome. The percentage of patients diagnosed as having the paranoid subtype was unrelated to treatment outcome.

Geographical location and the type of facility in which the studies were conducted both affected the results. Studies from non-Western countries (six from China and two from Israel) tended to produce higher effect sizes, while studies from Scandinavian countries and the United States and Canada yielded smaller effect sizes ($Q_B = 51.40$, $df = 4$, $p < 0.001$; table 3). This effect was statistically significant before and after adjustment for multiple testing. However, even after removing the studies from the non-Western and Scandinavian countries, the effect sizes from the remaining countries were still heterogeneous ($Q_B = 11.01$, $df = 2$, $p = 0.004$). Studies from the United States and Canada tended to produce lower effect sizes compared with studies from Great Britain and Continental Europe. When we removed studies from the United States and Canada, the heterogeneity test was no longer significant, indicating that studies from Great Britain and Continental Europe produced similar

results. Studies from VA hospitals produced smaller effect sizes (Q_B = 12.08, df = 5, p = 0.025; table 3). This effect was statistically significant only before adjustment for multiple testing and not after such adjustment. When studies from the VA hospital were excluded from the analysis of type of facility, the between-groups heterogeneity statistic was not statistically significant (Q_B = 2.25, df = 4, p = 0.690).

The impacts of specific treatment characteristics (i.e., modality and orientation of treatment) were examined separately and are reported in the next section. However, we also examined the effect of such general moderators as setting (inpatient vs. outpatient), duration of treatment, frequency of sessions, total number of sessions, and total number of treatment hours. Of these variables, duration of treatment had a statistically significant impact on the results (r = 0.48, df = 41, p = 0.0013), with and without adjustment for multiple testing. However, examination of the data revealed an outlier: A study of community care extending over 4 years by Madianos and Madianos (1992) yielded an estimated effect size of 1.6, as much as three SDs above the mean for the studies included in this analysis. After excluding this study, the correlation of duration of treatment with the effect size was no longer reliably larger than zero (r = 0.23, df = 40, p = 0.148). In addition, neither total number of treatment hours (r = 0.31, df = 39, p = 0.052) nor the total number of sessions (r = –0.16, df = 27, p = 0.416) had any reliable impact on the results.

Among the few therapist variables coded in our review, the number of therapists and therapists' experience did not have any significant impact on the results of the studies. The variable of therapists' allegiance did not have acceptable interrater reliability and, therefore, was not used.

Impact of Modality. The psychosocial treatments administered in these studies could be classified into six basic modalities: individual, group, family, milieu, occupational/recreational, and community care. There was a statistically significant difference between effect sizes for these modalities, as presented in table 5 (Q_B = 11.7, df = 5, $p < 0.05$). Studies reporting on the effects of group therapy produced the smallest effect sizes. When these studies were removed from the sample, there were no differences between estimates from the other five modalities (Q_B = 0.93, df = 4, p = 0.920).

Impact of Orientation. The psychosocial treatments in these studies could be classified into three broad theoretical orientations: behavioral, "verbal" therapies, and cognitive training. The first two classes are similar to the classes used in the general meta-analysis of Smith et al. (1980) for classifying types of psychotherapy. The verbal therapies, as defined in this classification, included a broad class of treatment approaches ranging from specific treatments such as gestalt therapy in a group format (Serok et al. 1984) and expressive insight-oriented psychotherapy administered in an individual format (Gunderson et al. 1984) to family therapies with a focus on expressed emotion (Barrowclough and Tarrier 1990).

The range of treatments in the broad theoretical class of behavioral treatments was also wide and included such

Table 5. Effect size estimates for the major modalities and orientations of treatment

Modality/orientation	d_+	95% CI	Q_W[1]	p
Modality				
Individual (10)	0.46	0.26 < δ < 0.66	20.79	0.014
Group (26)	0.25	0.14 < δ < 0.36	24.03	0.315
Family (12)	0.45	0.29 < δ < 0.60	16.12	0.137
Milieu (14)	0.47	0.35 < δ < 0.58	59.68	<0.001
Recreational/occupational (4)	0.41	0.08 < δ < 0.74	0.73	0.866
Community care (4)	0.45	0.29 < δ < 0.61	28.84	<0.001
Orientation				
Behavioral (13)	0.41	0.26 < δ < 0.56	14.77	0.254
Social skills training (2)	0.44	0.05 < δ < 0.83	5.76	0.016
Other behavioral programs (11)	0.41	0.25 < δ < 0.57	8.99	0.533
Verbal (46)	0.37	0.30 < δ < 0.44	143.57	<0.001
Psychodynamic psychotherapies (10)	0.27	0.11 < δ < 0.42	20.30	0.016
Expressed emotion reduction programs (7)	0.56	0.33 < δ < 0.79	2.76	0.838
Other verbal treatments (29)	0.38	0.30 < δ < 0.46	117.51	<0.001
Cognitive training programs (11)	0.41	0.20 < δ < 0.61	10.35	0.410

Note.—Numbers in parentheses indicate the number of studies. CI = confidence interval.

[1]Q_w is a measure of within-study heterogeneity. Statistically significant values indicate that the effect sizes from studies within a class are heterogeneous.

interventions as token economy administered in a milieu format (e.g., Paul and Lentz 1977), relaxation and systematic desensitization in an individual format (Weinman et al. 1972), social skills training, and other less commonly used techniques.

Among the studies in the behavioral and verbal therapies classes, we could identify five narrower orientations and calculate effect sizes for these subclasses as well: social skills training, other behavioral programs, psychodynamic psychotherapies, family treatments with a focus on expressed emotion, and other verbal therapies (table 5). There were no statistically significant differences between effect size estimates from the three broad orientations ($Q_B = 0.23$, $df = 2$, $p = 0.891$) or the five narrower ones ($Q_B = 4.23$, $df = 4$, $p = 0.376$).

Posttreatment Versus Followup. A previous meta-analysis of more than 60 studies suggested that effect sizes from psychotherapy at posttreatment do not differ significantly from effect sizes at followup (Nicholson and Berman 1983). However, this study was limited to the range of disorders traditionally labeled as neurotic and the authors entertained the possibility that, for other types of disorders, results at followup may differ from results obtained at posttreatment. We were able to examine this possibility for patients with schizophrenia using 10 studies that reported both posttreatment and followup results. Followup measurements were obtained from 1 week to 24 months after posttreatment measures, with a median of 12 months. The effect size estimates and 95 percent CIs for these 10 studies were $d_+ = 0.38$ ($0.32 < \delta < 0.44$) for posttreatment and $d_+ = 0.42$ ($0.24 < \delta < 0.59$) for followup. We used the method of generalized least squares suggested by Raudenbush et al. (1988) to model these 20 effect sizes as a common effect size versus two separate effect sizes. Neither model fit the data, reflecting a large proportion of variance in these effect sizes that is not explained by the distinction between posttreatment versus followup measures. However, comparison of these two models resulted in a nonsignificant chi-square difference test (χ^2 [1, $n = 20$] = 1.97, $p = 1.0$), suggesting that modeling followup separately from posttreatment does not add significantly to prediction of effect size variation. Hedges and Olkin (1985) have proposed a method for testing the homogeneity of correlated effect sizes (pp. 210–213). This method was applied separately to each of the 10 pairs of effect sizes. In nine cases, the test of homogeneity was not rejected, suggesting that posttreatment and followup effect sizes are not different. In the only study in which the test of homogeneity was rejected, the effect size for followup was significantly larger (Bellack et al. 1984). These findings suggest the enduring benefits of psychosocial treatments for patients with schizophrenia. The costly followup design may be used more judiciously in studies of patients with schizophrenia as well as patients with other diagnoses (Nicholson and Berman 1983).

Multiple Regression Analysis

Some of the variables that explained the between-study variance in the preceding univariate analyses may be confounded with each other. For example, consider the hypothetical case that some types of studies (e.g., group treatments or studies from Scandinavian countries) more commonly used samples of acute patients, who, in turn, tended to produce smaller effect sizes. In this hypothetical case, the lower effect size in these types of studies is attributable to a confounding factor: duration of illness. Multiple regression can untangle the effects of the different confounding variables by taking into account the impact of several variables at the same time and identifying the added contribution of each in the presence of others. (Univariate analysis, by its nature, is incapable of this task.) To this aim, we used WLS regression analysis, applying stepwise strategies suggested by Hedges (1994, pp. 297–298). We conducted two analyses. In the first, we consecutively entered into the model all the variables that were available for the 70 studies in the sample. In the second analysis, we added the variable of chronicity of illness, which was available for only 32 studies, to the predictor variables remaining in the final model of the first analysis. Because of the larger sample size, results of the first analysis are probably more robust; results of the second analysis should be interpreted with more caution.

For these analyses, the following variables were entered in order and retained or removed according to whether they explained any of the variation in effect size not accounted for by the previous variables: the number of outcome measures; the number of effect sizes set to zero; source and content of outcome measures (the variable of interest in this case was the number of outcome measures from each of the different sources and contents); publication date; sample size; allegiance of authors; basis of diagnosis; geographical location of the study; type of facility; and modality of treatment.

Only four variables emerged as significant at the end of the first analysis. The number of measures of disorganized behavior and employment were significantly associated with the effect size. Studies that used these measures reported lower effect sizes than studies that did not. Also, the effects of geographical location of the study (Scandinavian vs. other and non-Western vs. other) and the basis for diagnosis (objective criteria vs. no objective criteria) were significantly associated with the outcome in

the multiple regression analysis. This analysis yielded a multiple correlation of $R = 0.704$ and a Birge ratio (Hedges 1994, p. 298) of $R_B = 1.38$, suggesting that 62 percent of between-studies variation was explained by these variables given the within-studies sampling variance. It has been suggested that the Birge ratio may be better than squared multiple correlation as an estimate of the proportion of explained variance (Hedges 1994).

For the second analysis, chronicity of illness was entered after the sequence of variables described above, which again resulted in a significant multiple correlation ($R = 0.790$, $p < 0.01$, $R_B = 1.24$), indicating that 76 percent of the between-studies variance was explained by the predictor variables. Whereas in this analysis chronicity of illness was significantly associated with the effect size, the number of employment outcome measures and the basis for diagnosis were no longer significant predictors. The number of measures of disorganized behavior and the geographical location of the study showed significant association with the effect size in this analysis as well. Because of differences in sample size, direct comparison of the two models produced in these analyses was not possible.

In summary, two conclusions are supported by the results of these regression analyses. First, a large amount of the variation in the effect sizes from this heterogeneous group of studies is explainable by a small number of the variables chosen. Second, studies from Scandinavian countries and studies using measures of disorganized behavior and possibly measures of employment tend to produce smaller effect sizes. Studies from non-Western countries, studies with more chronic patients, and possibly studies using objective diagnostic criteria tend to produce larger effect sizes. The impact of measures of unemployment and objective diagnostic criteria is not supported with the same degree of confidence as the impact of other variables because these two variables did not contribute to the prediction of outcome in the second regression analysis when the variable of chronicity was added to the model.

Discussion

This study addressed important questions concerning the role of psychosocial treatments in management of patients suffering from schizophrenia. Although the utility of combining psychosocial and somatic interventions in the management of schizophrenia has gained general acceptance, no previous quantitative review has provided a comprehensive comparison of combined and somatic treatments to document this effect and provide an estimate of its magnitude. The present study addressed this need. Findings from this review demonstrated that patients who received psychosocial treatment in addition to somatic treatment scored, on average, 0.39 SDs higher (reflecting greater improvement) on measures of outcome than those who received only somatic treatment. An effect size of this magnitude implies that the average patient in the combined treatment group was more improved than 65 percent of the patients in the somatic treatment group. Using the binomial effect size display of Rosenthal and Rubin (1982), this effect size translates into success rates of 69 and 31 percent for combined and somatic treatments, respectively.

Underscoring the efficacy of the combination of psychosocial and somatic treatments, the findings from this review demonstrated that the combined treatments maintained their relative advantage over somatic treatments alone across a median followup period of 12 months. In 9 of 10 studies examining the durability of gains, there were no differences between posttreatment and followup effects, and in the remaining study, the effects at followup were larger than those at posttreatment. Moreover, our results revealed that combined interventions consistently produced lower relapse frequencies than somatic treatments alone. Fourteen studies reported such findings across a median period of 17 months. In those studies, combined treatments yielded relapse frequencies that were, on average, 20 percent lower than those produced by somatic treatments.

Our estimate of the beneficial effect of combining psychosocial with somatic treatments is comparable to the findings reported by Smith et al. (1980). In their overall analysis of the effects of drug therapy and drug plus psychosocial therapy, based on 566 effects from 112 studies, those investigators estimated the added effect of psychotherapy to be 0.31 SDs. However, that estimate was obtained from a heterogeneous set of studies that included patients with various diagnoses and various medication groups such as antidepressants, anxiolytics, and antipsychotics; our review was limited to patients with schizophrenia who were being treated with antipsychotics or, in a few cases, ECT. In their analysis of studies of schizophrenia patients, Smith et al. (1980) reported effect sizes of 0.495 for medication only and 0.802 for medication plus psychosocial treatment. The difference between these estimates, 0.31, is comparable to findings from their overall analysis and only slightly smaller than our estimate of the added benefit of psychosocial treatment. However, the Smith et al. (1980) estimate for combined treatment was based on only 29 effect sizes, and the estimate for medication was based on only 108 effect sizes. As noted previously, a significant limitation of the Smith et al. (1980) meta-analysis was the use of individual effect sizes as the units of analysis, a procedure that creates nonindependence in the data and arbitrarily weights studies by the number of outcome measures reported.

The only other general meta-analytic review of treatments for schizophrenia, the Quality Assurance Project (1984), reported estimates for the effects of the psychosocial treatments alone or in conjunction with medication. In that study, psychosocial treatments were grouped into two general classes: psychotherapy and social intervention. Whereas the effect of psychotherapy alone was estimated as zero, social intervention alone was estimated to yield an effect size of about 0.2. Further, the estimates for the combined treatments differed widely between these two forms of psychosocial intervention. The effect size for medication plus psychotherapy was estimated to be around 1.0, equal to the effect of medication alone. In contrast, the effect size for social treatment combined with medication was estimated to be around 2.0, twice as large as the effect of medication alone. Unfortunately, the definition of "psychotherapy" and the kinds of treatments included under that rubric in the Quality Assurance Project study were unclear. Further, the representativeness of the findings from the review were compromised: Only 26 studies were included, and the effect sizes were not calculated for all the measures used in each study.

In our meta-analysis, we not only attempted to determine the overall efficacy of combined psychosocial and somatic treatment relative to somatic treatment alone, but we also examined the impact of various moderator variables on the magnitude of the treatment effect. Both univariate and multivariate analyses were conducted for this purpose. We identified several significant characteristics associated with effect size in univariate analyses but not in multivariate analyses (e.g., sample size, publication date, allegiance of the authors). Such variables may be important moderators, but because they are correlated with other variables that have a stronger relationship with effect size, their impact is reduced in multivariate analyses. Of course, interpretation of these findings should be tempered by the recognition that the data are correlational and do not support strong causal inferences. The impact of certain general characteristics such as sample size and authors' allegiance have been discussed in previous meta-analyses (Smith et al. 1980; Robinson et al. 1990). In the following discussion, we focus on those findings that are unique for our meta-analysis and for this population.

Our analyses revealed that estimates of treatment effect varied as a function of the content of the outcome measures used, a finding of potential theoretical interest. Specifically, studies using measures of disorganized behavior and employment reported smaller effect sizes. Indeed, as shown in table 4, these two types of measures produced the lowest effect sizes among measures of various contents (0.13 and 0.22, respectively). Objective measures of adjustment, general symptom measures, anxiety/depression, positive symptoms, and cognitive functioning yielded modest effects (0.26 to 0.36), and larger effects were observed on measures of thought disorder, negative symptoms, compliance, and relapse (0.46 to 0.59). Several groups of researchers have proposed a three-factor structure for symptoms of schizophrenia: including negative symptoms, positive symptoms, and disorganization, with the last mostly defined by disorganized behavior (Andreasen et al. 1995; Arndt et al. 1995). Our finding of substantially lower effect size estimates for measures of disorganized behavior suggests that this dimension is not as responsive to psychosocial treatments as other dimensions, lending support to its delineation as a distinct dimension.

Also of theoretical interest is the finding that the effect size for negative symptoms was larger than that for positive symptoms, although the difference was not statistically significant in multivariate analyses. Studies of the effects of antipsychotic medications have consistently reported greater improvement in positive symptoms than in negative symptoms (e.g., Arndt et al. 1995). The considerable effect of combined therapy on the negative symptoms as demonstrated in our review would suggest that psychosocial treatments can not only augment the effects of medication, but also supplement them by improving symptoms less affected by conventional medication.

Of the various patient-related moderator factors coded in this study, chronicity of illness appeared to have the most substantial impact on the outcome. Studies including patients with more chronic illness produced larger effect sizes than studies with more acute patients, suggesting that psychosocial treatments were more beneficial in the chronic stages of illness. If so, such a pattern would differ from that observed for somatic treatments, which tend to be more effective in controlling symptoms of acute illness (Szymanski et al. 1996). This observation has important implications for treatment and general care of patients suffering from schizophrenia. Although most authors writing about the management of schizophrenia have stressed a combination of medication and psychosocial treatment, they have adopted a largely static, cross-sectional point of view. The possibility of an interaction between the chronicity of illness and the effectiveness of psychosocial treatments argues for a longitudinal and dynamic approach (Strauss 1989) in which different treatment priorities are adopted for different stages of illness.

An interesting finding from the present review was the effect of geographical location of the study. Studies from non-Western countries reported larger effect sizes, and among studies from Western countries, those from Scandinavian and North American countries reported smaller effect sizes than the others. In fact, studies from Scandinavian countries appeared to yield no reliable

effects because the CIs for their effect sizes included zero. The impact of geographical location (non-Western vs. other; Scandinavian vs. other) persisted in regression analyses.

This finding is open to more than one interpretation. For example, the observed differences may represent genuine differences in responsiveness of patients from these diverse cultural, ethnic, and geographical locations. Alternatively, the standard treatments provided in Scandinavian and North American countries may be more rigorous, limiting the possible impact of the added experimental treatment. This phenomenon is similar to the "ceiling effect" in which introduction of an additional intervention for a group that is already functioning at a high level would improve functioning only minimally. If, in fact, the standard treatments in the Scandinavian and North American countries are more rigorous, the difference between the experimental groups and the control groups in studies from these countries would be smaller than in those from other countries. On the other hand, if the standard treatment in the non-Western countries is poor, the impact of any added experimental treatment would be more prominent.

Out of the different settings, studies conducted in VA hospitals produced the lowest effect sizes in univariate analysis. Interpreting this finding is difficult. It is possible that higher standards of care for these patients in general leads to better functioning in both the experimental and comparison groups, and as a result, the experimental treatment can add little.

One of the puzzling findings from this review was the impact of diagnostic criteria on estimates of treatment effect. Studies in which diagnoses were based on objective criteria, whether narrow or broad, produced larger effect sizes than studies in which diagnoses were not based on objective criteria. Recent studies tended to use objective diagnostic criteria more often; however, entering publication date in multiple regression analysis did not eliminate the impact of diagnostic criteria on magnitude of treatment effect.

Finding a "best treatment" or a "treatment of choice" has been a persistent preoccupation of workers in the field of psychotherapy outcome research as well as in other clinical and medical fields. Identification of the most effective modality or theoretical orientation was one of the questions that motivated our study as well. Modality of treatment had significant effects in univariate analysis: group treatment produced smaller effect sizes compared with other modalities of treatment. However, this effect did not persist in multivariate regression analysis and may be a confound. Given the continued popularity of group treatments for this population, the effectiveness of these treatments compared with other modalities needs to be addressed in future primary studies.

The use of psychodynamically oriented therapies in the treatment of schizophrenia has been controversial. Some authors argue forcefully for psychodynamic treatment (e.g., Karon 1989); others describe such approaches as ineffective and possibly even harmful and call for a moratorium on their use (e.g., Mueser and Berenbaum 1990). The most exhaustive study comparing psychodynamic treatment with supportive therapy (Gunderson et al. 1984) yielded an effect size of –0.02, an effect not significantly different from zero. In our review, treatments with verbal and behavioral orientations had roughly equal effects, and within the verbal category, therapies based on various psychodynamic principles were not significantly less effective than verbal treatments based on other theoretical rationales. Thus, our review provides no evidence that psychodynamic therapies are harmful, nor does it suggest that psychodynamic treatments are superior to other interventions.

Given the broad categories of orientation and modality used in this review, as well as the substantial heterogeneity of effect sizes within the verbal category in particular (see table 5), the absence of statistically significant differences across theoretical orientations should be viewed with caution. For example, treatments focused on reducing expressed emotion in families produced an average effect size that was twice as large as the effect size for psychodynamic therapies (0.56 and 0.27, respectively). The failure to find statistically significant differences in this analysis could be attributed to low statistical power due to the small number of primary studies available or to the large variance within categories of orientation. Nevertheless, this pattern is at least suggestive of potential differences. As more primary studies become available, future meta-analyses may be able to detect statistically significant differences between treatments varying in theoretical orientation. In this regard, the finding in this and other meta-analyses (e.g., Berman et al. 1985; Robinson et al. 1990) of a significant relationship between allegiance of the researchers and effect size suggests that studies attempting direct comparisons between therapies of different theoretical orientation should be carried out as collaborative efforts, with adherents of each treatment being equally involved in the design and conduct of the investigations.

Limitations

As implied in the foregoing discussion, the conclusions drawn from this review should be tempered by consideration of the limitations of the meta-analytic methods and

our application of those methods to the available literature on the treatment of schizophrenia. One notable limitation of most meta-analyses, including the present one, concerns the nature of correlations examined: A large proportion of those correlations are between studies rather than within studies. In examining the relationship between the characteristics of patients in the studies and the estimates of treatment effect, correlations are calculated across studies using the average patient characteristics (or, for categorical variables, the percentage of patients exhibiting the characteristic) and the average effect size for each study. Because such correlations are between studies, it is difficult to control for the various confounds (i.e., other differences across studies that might explain the obtained correlations). For example, our finding of a substantial correlation between chronicity of illness and treatment effect across studies does not provide the same degree of confidence as evidence based on substantial correlation between chronicity of illness and response to treatment obtained within individual studies. Other patient, setting, and treatment characteristics (not all of which were measured or reported in individual studies) may account for the correlation we observed. It should be noted that this is not a limitation of meta-analytic methodology *per se*, but rather a limitation of the primary studies to which quantitative review techniques are applied. Thus, if a large enough number of primary studies had examined and reported such relationships within studies, the findings from those analyses could have been transformed into effect sizes (e.g., the within-study r between chronicity and treatment response) and integrated across studies. Moreover, some variables can be examined only across studies because examining their impact within studies is either not feasible (e.g., the impact of the time of the study) or not easy (e.g., the impact of geographical location).

A second limitation of our meta-analysis derives from our inability to examine interaction effects in the available primary studies. Like narrative reviews, meta-analyses are constrained by the number of primary studies testing a conceptual hypothesis, the specific variables examined in those studies, and the detail with which study characteristics are reported in published articles. In the present review, such constraints limited our analyses of moderator variables to tests of main effects. Small number of studies and incomplete reporting of information in the available research precluded an examination of important hypotheses, notably those involving interactions among setting, treatment, and patient characteristics. Investigation of the interaction between patient and treatment characteristics is of considerable clinical interest. It is conceivable that specific treatment modalities or orientations may be more efficacious in particular kinds of patients at certain stages of illness. However, the limited number of studies available makes the investigation of such interactions a prohibitive task. For example, studying the interaction of modality (six levels) and setting (three levels) would require a table with 18 cells and sufficient numbers of studies in each to ensure adequate statistical power for testing the interaction. Unfortunately, because of unequal numbers of studies representing each level, many cells in such a table would remain empty. In the present review, only 6 percent of studies were conducted in partial hospitalization settings, whereas two-thirds of the studies were conducted in inpatient settings. Similarly, about one-third of the studies examined group treatments, but only 6 percent investigated community-based interventions. Investigating the three-way interaction between these two variables and the treatment orientation would exacerbate the problem, requiring a table with 54 cells (for three levels of orientation) or 108 cells (for six levels of orientation).

Conclusion

In conclusion, although our quantitative review provides a useful summary of the available evidence and addresses some important questions regarding the efficacy of combined psychosocial and somatic interventions, it should be considered an interim report. In particular, we urge caution in drawing inferences regarding the impact of moderator variables on estimates of treatment effectiveness. Additional primary studies, more complete reporting of study characteristics and findings, and increased attention to interactions in those studies may permit future quantitative reviews to examine these important interactions and draw stronger inferences regarding the role of moderator variables.

Keeping in mind the limitations of this study, our findings may have implications for the practitioners working with patients suffering from schizophrenia. First of all, our results show that psychosocial treatments can play an important role in the comprehensive management of schizophrenia not only to augment the effects of medications, but also to supplement these effects in areas where conventional medications alone are less effective (e.g., negative symptoms). Second, there is some evidence that psychosocial interventions may be more effective in the more chronic stages of illness and, therefore, can play a more prominent role in the management of patients with chronic schizophrenia. Third, in view of the limited evidence for larger effects for other modalities over group treatments, it is advisable to choose treatments administered in an individual or family context over those administered in a group context. Also, we noted that family treatments focused on reduction of expressed emotion tended to have larger effect sizes than other orientations,

although this difference was not statistically significant. Obviously, in choosing between treatment options, a host of other factors (e.g., availability and cost) in addition to the relative efficacy of treatments needs to be taken into consideration.

Finally, our findings support the conclusion that patient characteristics and the common elements of therapy are more important as determinants of outcome than a particular modality or orientation. Better delineation of these patient characteristics and of common therapeutic elements is the task and challenge of future research.

References

American Psychiatric Association. *DSM–II: Diagnostic and Statistical Manual of Mental Disorders.* 2nd ed. Washington, DC: The Association, 1968.

American Psychiatric Association. *DSM–III: Diagnostic and Statistical Manual of Mental Disorders.* 3rd ed. Washington, DC: The Association, 1980.

American Psychiatric Association. *DSM–III–R: Diagnostic and Statistical Manual of Mental Disorders.* 3rd ed., revised. Washington, DC: The Association, 1987.

Andreasen, N.C.; Arndt, S.; Alliger, R.; Miller, D.; and Flaum, M. Symptoms of schizophrenia: Methods, meanings and mechanisms. *Archives of General Psychiatry,* 52:341–351, 1995.

Arieti, S. *Interpretation of Schizophrenia.* New York, NY: Brunner, 1955.

Arndt, S.; Andreasen, N.C.; Flaum, M.; Miller, D.; and Nopoulos, P. A longitudinal study of symptom dimensions in schizophrenia. *Archives of General Psychiatry,* 52:352–360, 1995.

Barrowclough, C., and Tarrier, N. Social functioning in schizophrenic patients: I. The effects of expressed emotion and family intervention. *Social Psychiatry and Psychiatric Epidemiology,* 25:125–129, 1990.

Bartko, J.J., and Carpenter, W.T., Jr. On the method and theory of reliability. *Journal of Nervous and Mental Disease,* 163:307–317, 1976.

Bartko, J.J.; Carpenter, W.T., Jr.; and McGlashan, T.H. Statistical issues in long-term followup studies. *Schizophrenia Bulletin,* 14(4): 575–587, 1988.

Beal, D.; Duckro, P.; Elias, J.; and Hecht, E. Increased verbal interaction via group techniques with regressed schizophrenics. *Psychological Reports,* 40:319–325, 1977.

Begg, C.B. Publication bias. In: Cooper, H., and Hedges, L.V., eds. *Handbook of Research Synthesis.* New York, NY: Russell Sage Foundation, 1994. pp. 399–409.

Bellack, A.S., and Mueser, K.T. Psychosocial treatment for schizophrenia. *Schizophrenia Bulletin,* 19(2):317–336, 1993.

Bellack, A.S.; Turner, S.M.; Hersen, M.; and Luber, R.F. An examination of the efficacy of social skills training for chronic schizophrenic patients. *Hospital and Community Psychiatry,* 35:1023–1028, 1984.

Benton, M.K., and Schroeder, H.E. Social skills training with schizophrenics: A meta-analytic evaluation. *Journal of Consulting and Clinical Psychology,* 58:741–747, 1990.

Berman, J.S.; Miller, R.C.; and Massman, P.J. Cognitive therapy versus systematic desensitization: Is one treatment superior? *Psychological Bulletin,* 97:451–461, 1985.

Carpenter, W.T., Jr., and Keith, S.J. Integrating treatments in schizophrenia. *Psychiatric Clinics of North America,* 9:153–164, 1986.

Cohen, J. *Statistical Power Analysis for the Behavioral Sciences.* New York, NY: Academic Press, 1977.

de Jesus Mari, J., and Streiner, D.L. An overview of family interventions and relapse on schizophrenia: Meta-analysis of research findings. *Psychological Medicine,* 24:565–578, 1994.

Falloon, I.R.H., and Liberman, R.P. Interactions between drug and psychosocial therapy in schizophrenia. *Schizophrenia Bulletin,* 9(4):543–554, 1983.

Feighner, J.P.; Robins, E.; Guze, S.B.; Woodruff, R.A.; Winokur, G.; and Munoz, R. Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry,* 26:57–63, 1972.

Fenichel, O. *The Psychoanalytic Theory of Neurosis.* New York, NY: Norton, 1945.

Fromm-Reichmann, F. *Principles of Intensive Psychotherapy.* Chicago, IL: University of Chicago Press, 1950.

Glass, G.V., McGaw, B.; and Smith, M.L. *Meta-Analysis in Social Research.* Beverly Hills, CA: Sage, 1981.

Grove, W.M., and Andreasen, N.C. Simultaneous tests of many hypotheses in exploratory research. *Journal of Nervous and Mental Disease,* 170:3–8, 1982.

Gunderson, J.G.; Frank, A.F.; Katz, H.M.; Vannicelli, M.L.; Frosch, J.P.; and Knapp, P.H. Effects of psychotherapy in schizophrenia: II. Comparative outcomes of two forms of treatment. *Schizophrenia Bulletin,* 10(4):564–598, 1984.

Heaton, R.K. *Wisconsin Card Sorting Test Manual.* Odessa, FL: Psychological Assessment Resources, 1981.

Hedges, L.V. Fixed effect models. In: Cooper, H., and Hedges, L.V., eds. *Handbook of Research Synthesis.*

New York, NY: Russell Sage Foundation, 1994. pp. 285–299.

Hedges, L.V., and Olkin, I. *Statistical Methods for Meta-Analysis.* San Diego, CA: Academic Press, 1985.

Hegarty, J.D.; Baldessarini, R.J.; Tohen, M.; Waternaux, C.; and Oepen, G. One hundred years of schizophrenia: A meta-analysis of the outcome literature. *American Journal of Psychiatry,* 151:1409–1416, 1994.

Karon, B.P. Psychotherapy versus medication for schizophrenia: Empirical comparisons. In: Fisher, S., and Greenberg, R.P., eds. *The Limits of Biological Treatments for Psychological Distress: Comparisons With Psychotherapy and Placebo.* Hillsdale, NJ: Erlbaum, 1989. pp. 105–150.

Landis, J.R., and Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics,* 33:159–174, 1977.

Lipsey, M.W., and Wilson, D.B. The efficacy of psychological, educational, and behavioral treatments: Confirmation from meta-analysis. *American Psychologist,* 48:1181–1209, 1993.

Maclennan, R.N. Interrater reliability with SPSS for Windows 5.0. *American Statistician,* 47:292–296, 1993.

Madianos, M.G., and Madianos, D. The effect of long-term community care on relapse and adjustment of persons with chronic schizophrenia. *International Journal of Mental Health,* 21:37–49, 1992.

Mueser, K., and Berenbaum, H. Psychodynamic treatment of schizophrenia: Is there a future? *Psychological Medicine,* 20:253–262, 1990.

Nicholson, R.A. and Berman, J.S. Is follow-up necessary in evaluating psychotherapy? *Psychological Bulletin,* 93:261–278, 1983.

Overall, J.E., and Gorham, D.R. The Brief Psychiatric Rating Scale. *Psychological Reports,* 10:799–812, 1962.

Paul, G.L. and Lentz, R.J. *Psychosocial Treatment of Chronic Mental Patients: Milieu Versus Social Learning Programs.* Cambridge, MA: Harvard University Press, 1977.

Quality Assurance Project. Treatment outlines for the management of schizophrenia. *Australian and New Zealand Journal of Psychiatry,* 18:19–38, 1984.

Raudenbush, S.W.; Becker, B.J.; and Kalanian, H. Modeling multivariate effect sizes. *Psychological Bulletin,* 103:111–120, 1988.

Robinson, L.A.; Berman, J.S.; and Neimeyer, R.A. Psychotherapy for the treatment of depression: A comprehensive review of controlled outcome research. *Psychological Bulletin,* 108:30–49, 1990.

Rosenthal, R., and Rubin, D.B. A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology,* 74:166–169, 1982.

Sarti, P., and Cournos, F. Medication and psychotherapy in the treatment of chronic schizophrenia. *Psychiatric Clinics of North America,* 13:215–229, 1990.

Schooler, N.R., and Keith, S.J. The clinical research base for the treatment of schizophrenia. *Psychopharmacology Bulletin,* 29(4):431–446, 1993.

Schwartz, B.J.; Cecil, A.; and Iqbal, N. Psychosocial treatments of schizophrenia. *Psychiatric Annals,* 23:216–221, 1993.

Serok, S.; Rabin, C.; and Spitz, Y. Intensive gestalt group therapy with schizophrenics. *International Journal of Group Psychotherapy,* 34:431–450, 1984.

Shadish, W.R.; Montgomery, L.M.; Wilson, P.; Wilson, M.R.; Bright, I.; and Okwumabua, T. Effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology,* 61:992–1002, 1993.

Smith, M.L.; Glass, G.V.; and Miller, T.I. *The Benefits of Psychotherapy.* Baltimore, MD: Johns Hopkins University Press, 1980.

Spitzer, R.L.; Endicott, J.; and Robins, E. Research Diagnostic Criteria: Rationale and reliability. *Archives of General Psychiatry,* 35:773–782, 1978.

Strauss, J.S. Subjective experiences of schizophrenia: Toward a new dynamic psychiatry–II. *Schizophrenia Bulletin,* 15(2):179–187, 1989.

Szymanski, S.R.; Cannon, T.D.; Gallacher, F.; Erwin, R.J.; and Gur, R.E. Course of treatment response in first-episode and chronic schizophrenia. *American Journal of Psychiatry,* 153:519–525, 1996.

Weinman, B.; Gelbart, P.; Wallace, M.; and Post, M. Inducing assertive behavior in chronic schizophrenics: A comparison of socioenvironmental, desensitization, and relaxation therapies. *Journal of Consulting and Clinical Psychology,* 39:246–252, 1972.

Wing, J.K.; Cooper, J.E.; and Sartorius, N. *The Measurement and Classification of Psychiatric Syndromes.* Cambridge, England: Cambridge University Press, 1974.

World Health Organization. *Manual of the International Classification of Diseases, Injuries and Causes of Death.* 8th ed. Geneva, Switzerland: The Organization, 1967.

World Health Organization. *Mental Disorders: Glossary and Guide to Their Classification in Accordance With the Ninth Revision of the International Classification of Diseases.* Geneva, Switzerland: The Organization, 1978.

The Authors

Ramin Mojtabai, M.D., Ph.D., is Postdoctoral Research Fellow, Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY. Robert A. Nicholson, Ph.D., is Associate Professor of Psychology, Department of Psychology, University of Tulsa, Tulsa, OK. Bruce N. Carpenter, Ph.D., is Associate Professor of Psychology, Department of Psychology, Brigham Young University, Provo, UT.

Announcement

The Child Psychiatry Branch of the National Institutes of Health is conducting a study on Childhood Onset Schizophrenia. Children 6 to 18 years of age who have been diagnosed with a psychotic disorder, who have had onset of psychotic symptoms by age 12, and whose family is compliant are being recruited. The study offers an intensive 3-week inpatient diagnostic evaluation for responders and nonresponders to current treatment. For nonresponders, this program offers a drug washout for diagnostic confirmation and a medication trial (approximately 3 months).

Benefits:

- Evaluation by a team that has seen more psychotic children than almost any other research facility in the country.
- Recommendations by psychiatrists, a social worker, nurses, teachers, and occupational and recreation therapists for future treatment.
- All treatment is free; housing and transportation are provided to those living at a distance.
- Opportunities for a drug-free washout for children who participate in the medication trial.

All referrals that meet the study criteria are welcomed.

For further information please contact:

Gale Germain
Building 10, Rm. 3N202
10 Center Drive, MSC 1600
Bethesda, MD 20892-1600
Telephone: 301-435-4504
Fax: 301-402-0296
E-mail: GEI@CU.NIH.GOV