Rehabilitation research methods for schizophrenia

Patrick W. Corrigan

Camarillo/UCLA Clinical Research Center for Schizophrenia and Psychiatric Rehabilitation, Camarillo, CA 93010, U.S.A.

(Received 19 April 1989, revised received 21 July 1989, accepted 25 July 1989)

Given the ever increasing complexity of rehabilitative models of schizophrenia, research methods testing these models must be rigorous and include a wide range of investigational strategies. This paper reviews four elements of rehabilitation research: definition of independent variables, selection of dependent variables, setting up the research design, and the development of conservative inferences from the data analysis. Methodological decisions must be made carefully prior to implementing a research protocol to assure the most valid conclusions when the study is complete.

Key words: Psychiatric rehabilitation; Research design

INTRODUCTION

Psychiatric rehabilitation describes an intervention strategy which facilitates patients' "physical, emotional, and intellectual skills needed to live, learn, and work in the community" (Anthony and Liberman, 1986, p. 542). Rehabilitation strategies have been dominated by environmental models which combine educational methods to facilitate the acquisition of prosocial skills and case management methods to augment the schizophrenics' limited support network. However, many chronic schizophrenics are unresponsive to psychiatric rehabilitation suggesting that researchers need to look beyond environmental models when generating hypotheses. Models immediately relevant to rehabilitation research include perspectives on symptomatology, cognition, social functioning, and psychophysiology.

Rehabilitation research methods test resulting hypotheses using descriptive or comparative formats. Descriptive investigation involves multifactorial approaches in which variables encompassed in biological and environmental perspectives are interrelated and reduced to better explanatory factors. Treatment strategies develop from these factors, or more commonly from intuitive understanding of change principles, and provide the main form of comparative research: treatment evaluation. Treatments can be evaluated using single case design, repeated measure group design, or comparison with alternative treatments. Careful application of clinical research methods strengthen the validity of conclusions drawn from rehabilitation investigations (Campbell and Stanley, 1963; Luborsky et al., 1975; Mosher and Keith, 1979; Glass et al., 1981). Rather than rehash these works in toto, this paper reviews methods specific to rehabilitation research of schizophrenia: definition of independent variables including subject variables and treatment protocol, selection of dependent variables from eight relevant domains, consideration of appropriate control groups in setting up the research design, and development of conservative inferences from the results. This paper serves as a primer for young researchers who pursue questions in this area. Moreover, review of these issues provides common standards to which senior investigators may refer in developing and reporting studies.

INDEPENDENT VARIABLES

Classification of schizophrenia

In order to conclude that a target treatment had beneficial outcome on subjects with schizophrenia.
independent assessment techniques must establish that the sample was in fact schizophrenic and not another psychopathology (Liberman et al., 1982). This can be a difficult task given the multitude of diagnostic criteria and is compounded by significantly different nosologies in Europe and the United States (see Table 1). Many European diagnostic systems are based on the presence of pathognomonic markers thereby producing relatively broad categories, e.g., patients who hallucinate are schizophrenic. American systems on the other hand, believe that the pathognomonic viewpoint does not distinguish disease-specific characteristics from pathological behavior that may correlate with chronicity or iatrogenicity. Criteria for diagnosis in the American system are narrower. As a result, research can be difficult to compare across the Atlantic, with American samples comprising a more homogeneous group than those drawn in the United Kingdom and on the continent.

Diagnostic systems differ in prognostic utility as well. Research has been unable to demonstrate that neither European nor American nosologies produce better predictions. Research Diagnostic Criteria, Carpenter's Flexible Criteria, and Langfeldt's Criteria were significantly correlated with outcome indicators while the New Haven Schizophrenia Index, Schneider's First Rank Symptoms, and the Catego System did not predict prognosis (Kendell et al., 1979). More recently, McGlashan (1988) found that Feighner's St. Louis Criteria was strong and the Research Diagnostic Criteria was moderate to strong in predictive validity. Once again, the New Haven Schizophrenia Index and Schneider's Criteria were weak prognosticators.

Frequent revision of the Diagnostic and Statistical Manual has affected the definition of psychopathology over time. DSM-II (APA, 1968) yielded more heterogeneous groups than DSM-III (APA, 1980). With the revision of DSM-III, criteria have changed again. A schizophrenic diagnosis on DSM-III-R (APA, 1988) requires meeting more symptom criteria, but no longer stresses deteriorated social functioning. With DSM-IV in the works, criteria will change one more time. While the ever increasing specification of schizophrenia and other pathologies is laudable, the ability to compare cohorts across different decades is hampered.

Diagnostic strategies differ among studies. Many investigations rely on one or two diagnostic systems popular in their geographic locale. More ambitious research define samples across several nosologies. This task is not facile, given the exclusivity of respective nosologies, and may require interviewing inordinate numbers of subjects to find a sufficiently large sample. Nevertheless, research programs must clearly specify the means by which patient characteristics were defined.

Several interviewing schedules have developed and are juxtaposed to their respective classification system in Table 1. While these schedules were developed with specific criteria in mind, most tend to glean broad information about the subject such that various nosologies can be applied. The Composite International Diagnostic Interview (CIDI) is a recent attempt by the World Health Organization and the Alcohol, Drug Abuse, and Mental Health Association to develop an instrument that will assess both ICD-10 and DSM-III-R psychopathologies (Robins et al., 1988). The CIDI includes items specific to schizophrenia and the affective psychoses from both the Diagnostic Interview Schedule and the Present State Examination. In addition, the CIDI includes assessment of substance abuse. The CIDI is a flexible instrument that can be administered by trained lay interviewers and can assess both present state and lifetime frames of reference.

Investigators need to describe the manner in which information was gathered to make the diagnosis. Diagnostic information can be gleaned from chart review, patient interview, or informant report. Chart review can be time efficient, but validity may be limited by the quality of progress notes. Moreover, investigators need to compare behavioral reports in the progress notes against diagnostic criteria and not rely on diagnoses assigned by clinicians. Patients can be fruitful historians, reporting private experiences that are difficult to observe but meet diagnostic signs. However, patient interviews may require several hours to complete. Moreover, information may be skewed by patient confusion and reactance. Informants (e.g., parents, siblings, other treatment staff) can validate information supplied by chart and patients.

Investigators should specify the diagnostic window in which criteria were assessed. In what time period must patients exhibit the symptom for it to
be considered current? Patient windows may vary from 2 weeks (‘in the past 2 weeks have you heard voices when nobody else is present?’) to a year. Diagnostic windows for interviews screening normal controls should extend for a lifetime (‘Have you ever believed you have special or magical powers?’). Lifetime windows assure investigators that normal comparison subjects do not contaminate results with subtle psychopathology.

Investigators should also report the diagnostician’s qualifications. Many interview instruments have specific probes and decision trees that can yield diagnoses by paraprofessionals. Other diagnostic systems require well trained individuals to establish psychopathological class. Assertions about the validity of diagnoses is supported by interrater reliability between two independent raters making diagnoses about the subject sample.

Schizophrenia has been subdivided into several classifications, e.g., paranoid versus nonparanoid patients (Magaro, 1980), positive versus negative syndrome (Andreasen, 1982a), process versus reactive presentation (Ullman and Giovannoni, 1964), and remitted versus deteriorating course (Luria and Meltzer, 1977). Investigators may wish to identify subclasses of treatment samples to determine the manner in which psychosocial intervention interacts with finer distinctions of schizophrenia. For example, research shows that paranoid and nonparanoid patients react differently to rehabilitation methods (Magaro et al., 1986). Similarly, subject demographics need to be described more fully. Research shows that schizophrenia varies across gender, race, and age (Klorman et al., 1977). Investigators may also include level of premorbid adjustment which encompasses educational and work achievements and peer relationships (see Table 2). At the minimum, specification of subject demographics allows comparison of treatment samples among independent studies.

Definition of treatment applications
For treatment comparison studies, independent assessment of the reliable implementation of treatment variables is necessary to conclude that outcome effects are due to intervention strategies and not extraneous constructs. Communication of this point is facilitated by reporting characteristics of the treatment protocol. Standardized protocols, in which treatment steps are prescribed in decision tree fashion, reduce the likelihood of error. Each step should allude to a well validated treatment strategy (e.g., token economy, modeling, role play, problem solving) or explicitly communicate the means by which an innovative strategy is carried out.

As remarked earlier, the efficacy of many rehabilitative strategies interact with schizophrenic sub-

---

**TABLE 1**

| Schizophrenic classification systems and instruments |
|----------------------------------|--------------------------------------------------|
| **Systems**                      | **Instruments**                                  |
| American                         |                                                  |
| (1) Diagnostic and Statistical   | Diagnostic Interview Survey                      |
| (2) Research Diagnostic Criteria | Schedule for Affective Disorders and             |
| (Spitzer et al., 1975)           | Schizophrenia (Spitzer and Endicott, 1978)      |
| (3) Diagno (Spitzer and Endicott, 1968) | Psychiatry Status Schedule (Spitzer et al., 1970) |
| (4) St. Louis Criteria (Feighner et al., 1972) |                                                  |
| (5) Carpenter’s Flexible Criteria (Carpenter et al., 1973) |                                                  |
| (6) New Haven Schizophrenia Index |                                                  |
| (Astrachan et al., 1972)         |                                                  |
| European                         |                                                  |
| (1) International Classification | Composite International Diagnostic               |
| Diseases-10 (WHO, 1989)          | Inventory (Robins et al., 1988)                  |
| (2) Catego (Wing et al., 1974)   | Present State Exam (Wing et al., 1974)           |
| (3) Schneider’s First Rank Symptoms (Schneider, 1959) |                                                  |
| (4) Crow’s types I and II (Crow, 1980) |                                                  |
| (5) Langfeldt’s Criteria (Langfeldt, 1960) |                                                  |
class. Therefore, protocols should specify the criteria by which subjects are included or excluded from participation in a study, i.e., demonstrate that a particular treatment is appropriate for a specific subsample (Carpenter, 1980). Some strategies are specific to phases of the disease, e.g., reduction of acute symptoms (neuroleptic medication) and facilitation of skills repertoire after chronic deterioration (psychoeducational methods). Other treatment modalities require a minimal level of functioning, e.g., insight-oriented talk therapy is not possible with a mute patient nor can acutely psychotic patients participate in highly stimulating rehabilitative programs. ‘Fair’ comparison studies require that subjects be appropriate for both target treatment and control group. Moreover, assigning heterogeneous subjects to the treatment group is likely to result in confounded negative findings while assignment of inappropriate subjects to the control treatment is likely to lead to type II error. Hence, investigators need to define conditions in which specific subject characteristics are indicated and methods used to assure that subjects adhered to the criteria for subgroup membership.

Investigators need to demonstrate that clinicians in all treatment conditions adhere to protocol prescriptions. In part, this requires that each clinician is competent in the studied treatment. Competence depends on scholastic achievement and specialized training in the strategies of the circumscribed treatment (Bernstein, 1982). Clinicians in each condition must attain similar competence so that outcome differences across conditions are not ascribed to unequal therapeutic skill. This goal can be achieved if clinicians are equally familiar with all treatment strategies and are rotated through every condition during the course of the experiment.

Independent methods can check staff fidelity to treatment protocols. Judges can observe treatment sessions and rate the presence or absence of key intervention skills that define the treatment condition. Treatment staff who do not utilize key skills are not adhering to the protocol. Comparisons across treatment groups can be adjusted by staff fidelity ratings in subsequent ANCOVAs. Similarly, when clinicians do not rotate through conditions, effects of individual treatment, staff can be evaluated by including therapists as a separate factor in multifactorial ANOVAs. Significant interactions among treatment and therapist factors suggests that clinical skill confounded outcome.

**DEPENDENT MEASURES**

Measures that are included in rehabilitation research must be sensitive to treatment effects. Occasionally, this requires construction of a tool specific to the hypotheses of the investigation. More commonly, however, investigators rely on already existing measures which have demonstrated reliability, validity, sensitivity, and non-reactivity. While the realm of variables relevant to schizophrenia is vast, current theoretical development makes eight domains of rehabilitation research salient: symptom level, cognitive abilities, psychophysiological responsivity, social adjustment and functioning, family functioning, quality of life, patient satisfaction, and inpatient days. Instruments measuring each domain are listed in Table 2.

The list of dependent measures is not exhaustive, but rather includes instruments validated in each domain. A review of each measure exceeds the goal of this paper. Instead, review articles germane to each domain are included in their respective discussions. Investigators reporting use of any dependent measure should heed the same rules outlined in the discussion of diagnostic classification, i.e., rater qualifications, interrater reliabilities, and means by which information was gathered.

**Symptom measures**

In conjunction with neuroleptic administration, rehabilitation programs are aimed at the reduction of positive and negative symptoms and the maintenance of remission. Both interview/observation instruments and self-report measures have been used to assess level of psychosis, although the thought-disordered patient is believed to less reliably self-report. Conversely, self-report measures are useful because they provide a phenomenological perspective on the psychoses not gathered in an interview. The Brief Psychiatric Rating Scale (Overall and Gorham, 1962) is frequently used for observation of symptoms while the MMPI and Symptom Checklist-90 are more common examples of self-report measures.
### TABLE 2

**Measures that are sensitive to rehabilitation related experimental manipulations**

<table>
<thead>
<tr>
<th>Variable domain</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall:</strong> interview</td>
<td>Brief Psychiatric Rating Scale (BPRS) (Overall and Gorham, 1962); Expanded BPRS (Lukoff et al., 1986); Psychotic Symptom Inventory (Silverstein and Harrow, 1987); Schizophrenia State Inventory (Grinker and Holzman, 1973)</td>
</tr>
<tr>
<td><strong>Overall:</strong> self-report</td>
<td>MMPI (Dahlstrom et al., 1975); SCL-90; Hopkins Symptom Checklist (Derogatis et al., 1974)</td>
</tr>
<tr>
<td><strong>Positive-negative symptoms</strong></td>
<td>Positive and Negative Symptoms Schedule (Kay et al., 1987); Schedule for Assessment of Negative Symptoms (Andreasen. 1982b); Schedule for Assessment of Positive Symptoms (Andreasen. 1984)</td>
</tr>
<tr>
<td><strong>Cognitive dysfunctions</strong></td>
<td>Bizarre idiosyncratic responses (Marengo et al., 1986); Thought, Language, Communication Scale (Andreasen, 1986); Thought Disordered Index (Solovay, et al., 1986); Exner’s Rorschach Criteria (Exner, et al., 1978)</td>
</tr>
<tr>
<td><strong>Abstract thought</strong></td>
<td>Comprehension Subtest, Wechsler Adult Intelligence Test-Revised (Wechsler, 1981); Goldstein-Scheerer Object Sorting Test (Goldstein and Scheerer, 1941); Gorham's (1956) Proverbs Test</td>
</tr>
<tr>
<td><strong>Information processing</strong></td>
<td>Continuous Performance Test (Kornetsky, 1972); Span of apprehension (Neale et al., 1969); Distractor digits (Rapaport, 1968); Visual masking (Balough and Merritt, 1987); Word recall tasks (Koh, 1978); Wisconsin Card Sorting Test (Goldberg et al., 1987); Word association (Broen, 1968); Reaction time (Nuechterlein, 1977a)</td>
</tr>
<tr>
<td><strong>Psychophysiological</strong></td>
<td>Smooth pursuit eye movement (Holzman and Levy, 1977); Electrodermal responsivity (Dawson &amp; Nuechterlein, 1984); EEG activity (Karson et al., 1988); Heart rate and blood pressure (Gruzuelier and Venables, 1975); Cortisol plasma levels (Spaulding et al., 1989)</td>
</tr>
<tr>
<td><strong>Social functioning</strong></td>
<td>Elgin Prognostic Scale (Wittman, 1941); Gittleman-Klein Scale (Gittleman-Klein and Klein, 1969); Goldstein Scale (Rodnick and Goldstein, 1974); Phillips' (1953) Prognostic Rating Scale; Abbreviated Phillips' Premorbid Adjustment Scale (Harris, 1975); Strauss and Carpenter's (1974) Prognostic Scales; Zigler and Phillips' (1961) Social Competence Scales</td>
</tr>
</tbody>
</table>

*Current adjustment: interviews*

<table>
<thead>
<tr>
<th>Variable domain</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavioral observations</strong></td>
<td>Katz Adjustment Scale (Katz and Lyerly, 1963); Personal Adjustment and Role Skills Scale (Ellsworth et al., 1968); Global Assessment Scale (Spitzer et al., 1976); Social Adjustment Scale (Weissman et al., 1971); Independent Living Skills Survey (Wallace, 1982); Nurses Observation Scale for Inpatient Evaluation (Honigfeld et al., 1966)</td>
</tr>
</tbody>
</table>

*Behavior self-report*

<table>
<thead>
<tr>
<th>Variable domain</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavior assertiveness</strong></td>
<td>Behavior Assertiveness Test (Eisler et al., 1975); Confederate's Test (Goldsmith and McFall, 1975); Means-End problem solving (Platt and Spivack, 1975); Simulated Social Interaction Test (Curran, 1982)</td>
</tr>
<tr>
<td><strong>Fear of negative evaluation</strong></td>
<td>Fear of negative evaluation (Watson and Friend, 1969); Social Avoidance and Distress Scale (Watson and Friend, 1969); Rathus (1963) Assertiveness Scale; Wolpe Lazarus (1966) Assertiveness Scale</td>
</tr>
</tbody>
</table>

*(Table 2 continued on next page)*
<table>
<thead>
<tr>
<th>Variable domain</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Life</td>
<td>Quality of Life Scale (Heinrichs et al., 1984); Quality of Life Checklist (Malm et al., 1981); Quality of Life Interview (Lehman, 1988)</td>
</tr>
<tr>
<td>Family Functioning</td>
<td>Camberwell Family Interview (Brown and Rutter, 1966); Abbreviated CFI (Vaughn and Leff, 1976); The Five Minute Speech Sample (Snyder and Liberman, 1981); Patient Rejection Scale (Kreisman et al., 1979); The Knowledge Interview (Berkowitz et al., 1981)</td>
</tr>
<tr>
<td>Communication deviance</td>
<td>Rorschach (Singer, 1968); Thematic Apperception Test (Jones, 1977)</td>
</tr>
<tr>
<td>Impact on family:</td>
<td>Need instrument development</td>
</tr>
<tr>
<td>Patient satisfaction</td>
<td>Perception of Ward Scale (Ellsworth and Maroney, 1972); Ward Evaluation Scale (Rice et al., 1963); Characteristics of Treatment Environment Scale (Allon et al., 1973); Comprehensive Consumer Satisfaction Scale (Essex et al., 1981); Ward Atmosphere Scale (Moos and Hout, 1968)</td>
</tr>
</tbody>
</table>

Number of days in hospital

Positive and negative symptoms describe opposite ends of a 'schizophrenic continuum'. These symptom clusters may have implications for etiology and treatment. Positive symptoms are commonly associated with the acting out schizophrenic and include hallucinations, inappropriate affect, delusions, bizarre behaviors, and conceptual disorganization (Andreasen, 1982a; Crow, 1985). Negative symptoms are characteristic of the withdrawn nontalkative patient with low energy and flat affect. Several interview-based instruments have been developed to measure these syndromes including the Schedule for Assessment of Negative Symptoms (Andreasen and Olsen, 1982) and its companion Schedule for Assessment of Positive Symptoms (Andreasen, 1984).

Cognitive measures
Three different approaches describe cognitive disorders and provide dependent measures sensitive to rehabilitation outcome. The classic view of cognitive deficit measures thought disorder in terms of pathologic verbal categories, e.g., impoverished speech, loose associations, flight of ideas, word salad, perseverations, and autistic logic. Andreasen (1986) developed the Scale for the Assessment of Thought, Language, and Communication to measure specific manifestations of this phenomenon. In addition, scoring systems have emerged to identify pathologic verbalizations on inflext tests (Exner et al., 1979; Johnston and Holzman, 1979).

Others viewed thought disorder as the schizophrenic's inability to manage abstract ideas (Harlow et al., 1986). This deficit is commonly seen as an inability to comprehend social rules or to explain the meaning of proverbs. While abstract cognition is diminished, schizophrenics show fairly intact concrete thinking skills.

The information processing model has attempted to locate thought disorder in specific stages or processes of cognition. Schizophrenics have shown below normal scores on tasks that assess attention (Nuechterlein, 1977), iconic memory (Sacuzzo, 1986), recall (Koh, 1978), and response selection (Broen, 1968). Generalized deficit models account for multiple processing limits based on an overall deficiency; examples include limited information processing capacity (Nuechterlein and Dawson, 1984a) and chronic hyperarousal (Gjerde, 1983). Measures of processing deficits are more typical of cognitive lab instruments. The Continuous Performance Test, Span of Apprehension Test, and Distractor Digit Span Test have been frequently used in schizophrenia research and meet the criteria for vulnerability markers of the disease (Nuechterlein and Dawson, 1984b).

Psychophysiological measures
Several indices are proven psychophysiological measures including smooth pursuit eye movement (Holzman and Levy, 1977), EEG alpha activity (Karson et al., 1988), electrodermal responsivity
(Dawson and Nuechterlein, 1984), plasma cortisol levels (Spaulding et al., 1989), heart rate, and blood pressure (Gruzelier and Venables, 1975). The relevance of psychophysiological and cognitive measures to rehabilitation research may not be readily apparent. The stress-vulnerability model of schizophrenia provides a synergistic view of biological and rehabilitation models (Nuechterlein and Dawson, 1984b; Zubin and Spring, 1977), in which psychophysiological and cognitive vulnerabilities emerge in subclinical fashion during adolescence or young adulthood. Psychophysiological deficits point to an autonomic dysfunction which exacerbate patients' sensitivity to stress (Dawson and Nuechterlein, 1984) while cognitive deficits prevent sufficient acquisition of coping skills further exposing the schizophrenic to stress. Relationships of cognitive/psychophysiological variables with social deficits have only recently been investigated, e.g., schizophrenics with significant fluctuations in autonomic indicators belong to families where parents were overinvolved or hypercritical (Tarrier et al., 1979, 1988; Sturgeon et al., 1984; Nuechterlein et al., 1989).

**Measures of social functioning**

Given its dependence on environmental models, rehabilitation research is commonly associated with measurement of social functioning, at premorbid, acute, or remitted/deteriorated stages. Strauss and Carpenter (1972, 1974) have argued that premorbid adjustment is highly prognostic, such that schizophrenics who have poor interpersonal and work histories are more likely to show a chronic course. Kokes et al. (1977) summarized interview measures which assess this construct. As discussed earlier, premorbid variables may be used as independent variables in treatment studies. Premorbid indicators may also serve as descriptors in multivariate research.

Morbidity assessment shows that social functioning covaries with acute and remitted stages. However, a subsample of schizophrenics demonstrates a chronic deteriorated course where symptoms never fully remit and functioning remains below premorbid levels. Assessment of morbid social adjustment levels is frequently used as outcome measures in rehabilitation research (Avison and Speechley, 1987; McGlashan, 1988). The instruments assess subject's functioning in work, social, and recreational spheres (Weissman, 1975; Weissman et al., 1981; Wallace, 1986). In addition, these instruments have formats in which significant others report patient progress.

Behavior assessment of social and coping skill offers an independent determination of individual social competence (Cone and Hawkins, 1977; Goldfried and Linehan, 1977). Behavior assessment typically involves role play formats in which subjects are instructed to accomplish instrumental or affectional goals with a confederate. Role plays are videotaped and rated in terms of criteria that assess micro- and molar-interpersonal skills. Though less valid, more time efficient self-report measures exist to assess level of interpersonal skills as well. Investigators question the validity of role play and self-report assessments, when compared with naturalistic observations, and argue that measures should be based on observation of real life behaviors (Curran, 1978; Bellack et al., 1979).

Paul (Paul and Lentz, 1977; Paul, 1987) developed the Time Sampling Behavior Checklist as a naturalistic observation strategy in which a range of targeted behavior is sampled several times each day to develop a frequency profile.

**Family measures**

Research has shown that family expressed emotion (EE, i.e., hypercriticality, overinvolvement) and communication deviance (CD) may exacerbate the schizophrenic family member's stress level and increase the likelihood of relapse (Vaugh and Leff, 1976). Several measures have developed to assess family EE with the Camberwell Family Interview being the benchmark by which other measures are evaluated (Brown and Rutter, 1966; Rutter and Brown, 1966). The CFI is time costly however, requiring more than 1 h to complete. Others have sought a shorter means for assessing EE (Snyder and Liberman, 1981). CD is typically measured by application of a complicated scoring system to the Rorschach (Singer, 1968) and the Thematic Apperception Test (Jones, 1977). Use of family-functioning variables have become more prevalent with the advent of behavior family management approaches to schizophrenia (Goldstein, 1981; Falloon et al., 1984; Anderson et al., 1986).

In addition to evaluating family impact on patients, dependent measures can determine the influence schizophrenia had on the family. Re-
ferred to as 'family burden', investigators in this area focus on the problems mental illness impose on the family, both as a set of individuals and as a system (Gubman et al., 1987; Tessler et al., 1987). Family problems include financial burden, social ostracism, and emotional impact as a result of prolonged concern about the patient. Unfortunately, instruments measuring this construct are not well developed.

Quality of life
In addition to diminishing symptoms and increasing social functioning, rehabilitation research needs to assess treatment effect on patients' quality of life (Bigelow et al., 1982). This is an increasingly popular, though somewhat elusive, construct defined as patient satisfaction with their attained goals and repertoire of skills. Thorough measurement of the construct should assess medical care, human relationship, work, communication and transportation, knowledge, and spiritual experience (Malm et al., 1981). The Quality of Life Scale (QLS) is the best developed measure with good interrater reliability and validation based on three factor analyses (Heinrichs et al., 1984). However, the authors have yet to publish a construct validation of the instrument, correlating the QLS with other measures. In addition to supporting the validity of the QLS, a construct validation will likely add definition to the quality of life concept.

Patient satisfaction measures
Patient satisfaction is a frequently neglected variable measuring patient attitudes on aspects of the treatment, the clinicians carrying out the treatment, and the clinic's physical surroundings (Weinstein, 1979; Le Bow, 1982, 1983; Nguyen et al., 1983). Measures of this construct include both self-report and interview formats, though interview measures may inflate satisfaction reports because of the demand characteristics of the interaction (LeVois et al., 1981). Interestingly, schizophrenics do not necessarily rate inpatient settings as less satisfactory than community clinics. Rather, each setting has components that are satisfying and displeasing to patients (Corrigan. 1989).

Hospitalization days
The number of inpatient days is frequently used as a measure of relapse and deteriorated functioning. Better inpatient rehabilitation programs should be able to decrease the number of days requiring hospital care. Furthermore, inpatients days are a necessary index to determine cost-benefit ratios, especially if these analyses involve comparison of hospital versus community treatment. However, conclusions based on change in hospital days must be made judiciously. The schizophrenics' hospital stay may not only reflect the course of disease and remission, but also agendas of the patients' community, family, and mental health treatment team who believe hospitals offer patients asylum from the real world or who desire a respite from the patients' maladaptive community behavior.

Selection of dependent measures
From which of eight domains of dependent measures do investigators select, when they set up a rehabilitation study? Many previous descriptive and comparative studies have focused on one or two domains, discussing specific findings in terms of these variables alone. However, limiting measurement to a few variables hampers the range of conclusions that may be drawn. Multidimensional research is necessary to evaluate overall treatment impact and facilitates examination of interactions of measurement domains.

Despite the advantages of multidimensional research, there exists a cost to including many measures in a study. Each measure requires training of raters and establishment of rater reliability. Administering and scoring the tests can require much time as well. Besides the costs, investigators need to beware of shot gun approaches in which small samples are administered many instruments. Significant results in these situations may represent a chance statistical finding, instead of a veridical relationship. Hence, variable selection is guided by a cost-benefit ratio; investigators must balance the time costs and possible false positives that infiltrate studies against the greater flexibility in addressing and answering rehabilitation questions.

RESEARCH DESIGN
Several parameters specifically affect rehabilitation designs including the number of comparison groups, the number of repeated measures on
groups, and matching subjects into pairs and subsequent random assignment (Luborsky et al., 1975). Two elements bear further discussion in research on schizophrenia: definition of control groups in comparison studies and specification of concurrent interventions in treatment comparison investigations.

When attempting to identify pathognomonic characteristics, 'normal' subjects make bad control groups. Schizophrenics do almost everything worse than normal individuals because of the generalized deficit that arises from having a chronic disease treated over a long course. Other thought disordered groups like manics make better control groups (Hoffman, 1987; Solovay et al., 1987). However, normal subjects may be useful in comparative studies for establishing treatment goals. Normal subjects provide useful data concerning 'typical' functioning and provide guidance for assessing patients' progress in terms of these norms.

Zubin and Spring (1977) defined a deficit as a unique disease vulnerability—i.e., a disease characteristic associated with relapse—by comparison of index measures which not only discriminate schizophrenics from other psychotic samples, but also distinguish remitted and acute patients, and disease risk individuals from same age controls. Maher (1965) described the differential deficit method as a second means of identifying disease characteristics. Schizophrenics who demonstrate significant decrements on task B, when both tasks are equated a priori for mean item difficulty and error variance (Chapman and Chapman, 1973, 1978) demonstrate a disease-related deficit in the construct measured by task A. For example, if schizophrenics perform worse on items measuring social information receiving skills (e.g., 'What did you see happen in that situation?') compared to items measuring processing of social data (e.g., 'What things could you do about the social event?') (Wallace, 1982), then receiving skill deficits are more likely characteristic of the disease.

Investigation write-ups of designs should also specify concurrent treatments. The patients' medication regimen should be described. Dose levels across different medications need to be converted to chlorpromazine equivalents (Davis, 1976). Strategies for changing doses as a result of symptom exacerbation should be defined as well, e.g., a criterion increase in BPRS scores (Lukoff et al., 1986). In addition, other aspects of concurrent treatment need to be described including status of residential placement, case management, and physical health care. Concurrent treatment should be equivalent across conditions in comparison designs. In this way, one can more certainly conclude that effects are due to targeted treatments and not extraneous interventions.

In addition to comparison treatment strategies, interventions can be studied using single case design (Barlow and Hersen, 1984). Single case investigations offer several advantages for pilot research. Preliminary hypotheses can be studied with very few subjects. Manipulation of treatment components can demonstrate their relative efficacy. Single subject methods may help to bring more practitioners into the research arena, thereby facilitating clinician-researcher dialogue.

CONSERVATIVE INFERENCEs

The ultimate purpose of rehabilitation research is to generate conclusions that will expand the knowledge base regarding schizophrenia and its treatment. The breadth and strength of conclusions can not exceed the power of the design or statistic. Analyses differ in terms of power efficiency, i.e., the test will reject the null hypothesis when the hypothesis is, in fact, false. Parametric tests have greater power thereby providing the strongest conclusions. However, parametric tests require adherence to more stringent assumptions: an interval or ratio scale of measurement, random assignment of subjects, normal distribution of dependent variables, and homoscedasticity. Nonparametric tests are more likely to result in type II error but can be applied to nominal or ordinal measurements or data sets that are not normally distributed. Meeting statistical requirements depends on a priori decisions about independent and dependent variables and design.

Similarly, conclusions should not violate the rules of deductive and inductive logic on which the scientific method rests. Unfortunately, these rules are frequently forgotten resulting in non sequiturs and overgeneralization. In comparative investigations, inferential limits are frequently ignored lead-
ing to overzealous statements regarding treatment efficacy and inappropriate assertions about treatment generalization to settings and with populations not evaluated in the study. Similarly, in multivariate research, correlational data are asserted to describe causal relationships rather than descriptive associations. Errors include mistaking relationships between extraneous variables as representative of real life associations. Finally, the need for replication is frequently ignored, perhaps in part because researchers do not find this work sufficiently reinforcing.

Investigations adhering to these rules will greatly expand the body of valid statements regarding psychiatric rehabilitation and schizophrenia.

ACKNOWLEDGEMENTS

Patrick Corrigan, Psy. D. is a postdoctoral scholar with the UCLA Department of Psychology.

The author acknowledges Robert Liberman and Charles Wallace for overall guidance.

This paper was made possible in part by an NIMH grant to the Clinical Research Center for Schizophrenia and Psychiatric Rehabilitation (MH30911, R.P. Liberman, Principal Investigator).

REFERENCES


American Psychiatric Association (1968) Diagnostic and Statistical Manual of Mental Disorders, 2nd edn. APA, Washington, DC.


American Psychiatric Association (1988) Diagnostic and Statistical Manual of Mental Disorders, 3rd revised edn. APA, Washington, DC.


paper on the validity of the role play test. Behav. Ther. 9, 462–468.
Gruzelier, J.H. and Venables, P.H. (1975) Evidence of high and low levels of physiological arousal in schizophrenia. Psychophysiology 12, 66–73.
Harris, J.G. (1975) An abbreviated form of the Phillips’ rating scale of premorbid adjustment in schizophrenia. J. Abnorm. Psychol. 84, 129–137.
Kornetsky, C. (1972) The use of a simple test of attention as a


