The Progress Evaluation Scales (PES) provide an efficient measuring device for evaluating current functioning, setting treatment goals, and assessing change over time in clinically relevant aspects of personal, social, and community adjustment. The PES can be completed by patients, significant others, and therapists, making it possible to obtain various points of view of the outcome of mental health services. This article describes the seven domains measured by the PES and the underlying dimensions they were designed to tap, and presents the generalizability, validity, and usefulness of the scales as applied to an adult *mental health center population.*

MEASURING PROGRAM OUTCOME

The Progress Evaluation Scales

DAVID IHILEVICH
Shiawassee County Community Mental Health Center
GOLDINE C. GLESER
University of Cincinnati
GORDON W. GRITTER
Michigan State University
LENORE J. KROMAN
University of Michigan
ANDREW S. WATSON
University of Michigan

With the advent of the deinstitutionalization over the past two decades and the massive shift of care to local communities, a growing clamor has developed for satisfactory methods to evaluate the quality of programs and services rendered at the community level. While such needs have been expressed from time to time by professionals and

AUTHORS' NOTE: *Financial assistance for this project was received for two years from a 314(d) Federal Grant and for six years from Michigan's Department of Mental Health.*

EVALUATION REVIEW, Vol. 5, No. 4, August 1981 451-477

agency administrators who wished to assess the efficiency and effectiveness of their efforts, more recently Congress, state legislatures, and local Mental Health Boards have begun requiring an explicit commitment to an ongoing process of evaluation of mental health services as a condition for continual social, political, and financial support.

In order to evaluate mental health services in a community, it is necessary to measure the outcome of interventions for clients treated in a number of diverse programs. Such measurement, however, has proven to be extremely complex, expensive, and time-consuming. A number of reviewers examining the state of the art have concluded that "program evaluation is still clearly in its infancy as a field of study" (Perloff et al., 1976: 586). The major difficulties noted in the literature on evaluation are: (a) a lack of consensus about stated or implied goals of various intervention techniques; (b) disagreement on the methods to be used to determine outcome; (c) uncertainty about the desirability and stability of observed change; and (d) unavailability of instruments for measuring outcome which are reliable, valid, and relevant for clinical and programmatic decision-making, yet are sufficiently broad to be applicable to the great variety of programs and clients served by mental health centers in the community.

The Progress Evaluation Scales were developed in an effort to overcome some of the difficulties noted in other evaluation approaches. Our purpose was to design an instrument that was (a) clinically and socially relevant; (b) administratively unobtrusive to the therapeutic process; (c) easy to implement in a busy, multidisciplinary agency; (d) economical for use on a continuing basis; (e) suitable for manual and computer data analysis; and (f) useful for clinical, programmatic, and policy decision-making at the clinic or mental health center level.

After two years of deliberation and preliminary testing, a consensus developed among the members of the research team, and the staff of the mental health center where the project was being developed, that the new instruments should be free from professional jargon so that they could be usable by clients, significant members of their family, and the assigned clinicians; they should enable respondents to select goals for various time intervals, as well as indicate current functioning levels;

We wish to thank the staff and the clients of the Shiawassee County Community Health Center for their extensive contributions toward the development of these scales. The material presented in this article is in the public domain. Requests for reprints should be sent to David Ihilevich, Shiawassee County Community Mental Health Center, 826 West King Street, P.O. Box 479, Owosso, MI 48867.

finally, they should be brief and easily understood so as not to take more than a few minutes to fill out.

Our planning and conceptualization were greatly influenced by recent advances in assessment of behavior change. The pioneering work of Kiresuk and Sherman (1968) on Goal Attainment Scaling has been particularly influential. They developed a system of translating traditional clinical terminology into behavioral descriptions which lend themselves to quantitative analysis. In addition, they conceived of an approach for establishing "entry status" at the beginning of treatment and of forecasting favorable and unfavorable outcomes over specified time intervals.

Another important influence on our thinking was the work of Ciarlo and Reihman (1974). They introduced the idea that goals of therapy and outcome of treatment should be determined by and measured against empirically derived norms which are representative of the actual functioning levels of people in the communities where the services are rendered.

DESCRIPTION

The Progress Evaluation Scales (PES) are made up of seven scales, each consisting of five levels, with the characteristics of each level described. For statistical purposes, the five points in each scale have been assigned a value of 1 to 5, from the most pathological to the healthiest levels of functioning observed in the community. All seven scales are printed on a single page for ease of administration and handling. Four slightly different versions of the scales are available to accommodate children (6-12), adolescents (13-17), and adult mental health clients, and the developmentally disabled.

The seven dimensions of the PES were chosen to represent the major areas in which health and psychopathology reveal themselves. This selection was influenced both by classical psychoanalytic theory and by other, more recent studies of positive mental health. Two general propositions underlie the psychoanalytic conception of mental health: first, health and psychopathology are end points of a single continuum, "no sharp line can be drawn [between them]... our conception of 'disease' is purely . . . a matter of degree" (Freud, 1909: 286); second, the healthy end of this continuum is characterized by the presence of a "capacity for work and enjoyment" (Fenichel, 1945: 581). In this citation, Fenichel

summarized Freud's position on mental health; that this position was considered relevant to the clinical domain is evident from Freud's advice to his fellow practitioners to consider it a therapeutic achievement "to win back part of the capacity for work and enjoyment" (Freud, 1912: 332). This psychoanalytic conception has been complemented more recently by a number of penetrating studies of positive mental health (Jahoda, 1958; Offer and Sabshin, 1966; Soddy and Ahrenfeldt, 1967). Whereas none of these studies have taken issue with the psychoanalytic propositions mentioned above, they have clarified the importance of additional factors such as meaningful personal relationships, positive self-regard, effective coping capacity, and modulated expression of affect, as essential ingredients for satisfactory personal, social, and familial adjustment.

As will be seen below, our selection and scaling of the PES dimensions reflect the notion of a single continuum for health and psychopathology, as well as the consensus that has emerged in the professional literature concerning the areas that should be scrutinized for the presence of evidence of positive mental health and psychopathology. Our strategy in constructing the scales was designed to measure change toward healthy adaptive functioning rather than relief of symptoms of personality reconstruction. This approach was selected since it appeared to approximate best the mandate of community mental health services. Descriptions of the scale points were developed with the help of the clinical staff and suggestions from numerous patients, as well as by empirical studies to determine score distribution, reliability, and relationship among the various scales.

The seven scales and the underlying dimensions they represent are as follows:

Family Interaction. This scale measures the dimension of "dependence-independence-interdependence" in one's relationship with other family members. The lowest level of functioning describes extreme dependent behavior where one needs help with such basic needs as eating and dressing; in the middle range one makes own plans and decisions but without necessarily considering the needs of other family members; at the highest level of functioning one plans and acts in such a manner that one's own needs, as well as needs of others in the family, are taken into account.

Occupation (School-Job-Homemaking). This scale taps a person's level of functioning in the "primary occupational role." At the lowest level of functioning, the person is unable to hold a job, care for home, or

go to school; at the highest level of functioning on this dimension, the person holds a regular job, attends classes, or carries out homemaker tasks (or some combination of these) with little or no difficulty.

Getting Along With Others. This scale was designed to tap the dimension of "socialization." The person's ability to establish and maintain satisfying relationships outside the family circle is a reflection of the degree to which this kind of socialization has satisfactorily occurred. The lowest level is characterized by a person who is always fighting, is destructive, or is alone; the highest level characterizes a person who gets along with others most of the time, and has close, regular friends. (Note that this scale does not attempt to differentiate the hostile person from the recluse; both are considered equally unsocialized.)

Feelings and Mood. This scale taps the level of "affective modulation" as indicated by the degree to which feelings are flexibly expressed and adaptively integrated into overall personality functioning. At the least satisfactory level of integration of affect, the person always feels nervous, depressed, angry, or bitter, or feels no emotions at all; the highest level of satisfactory affect integration is expressed in being in a good mood most of the time, as well as being able to be as happy, sad, or angry as the situation calls for. (Note that while the nature of affect at various levels is described, it is the persistence of certain affects which is measured.)

Use of Free Time. This scale assesses the degree to which "sublimatory processes" have satisfactorily evolved by indicating how free or constricted a person is in using inner and outer resources for play and enjoyment. The lowest level of functioning on this dimension is expressed in a person's almost total lack of interest in recreational activities of hobbies; the highest level of functioning is expressed in a person's participation in as well as ability to create a variety of own recreational activities and hobbies.

Problems. This scale taps the "coping capacity" the person can bring to bear on life's circumstances. At the lowest level of functioning the person is unable to handle even mild problems, hence, severe difficulties are experienced most of the time; at the highest level of functioning the individual is able to handle well even severe problems, therefore, the person is described as having only occasionally mild problems. (*Severe* refers to problems which are "incapacitating" in important areas, such as homemaking, work, sex, communication, or parenting: *Moderate* refers to problems which are "impairing" one's efficiency and/or effectiveness, but which are not totally incapacitating; *Mild* refers to prob-

lems which are "annoying or inconveniencing" but which do not incapacitate or interfere with a person's functioning in important areas in their lives.)

Attitude Toward Self. This scale assesses the dimension of "self-esteem." The lowest level score is characterized by having a negative attitude toward self most of the time; the middle level is described as having an almost equal positive and negative attitude toward self; the highest level is characterized as having a positive attitude toward self most of the time.

To use the scales (see Appendix A), therapists, clients and/or significant others *independently* indicate the initial status of clients by selecting the one item in each scale that describes best their current functioning level. "Current functioning" is defined as typical behavior and experience during two weeks preceding the evaluation interview.

After current status is indicated on one sheet for the seven scales of the PES, goals are set on a different sheet, again independently, by therapists, clients and/or significant others. For adults who can read at the level of comprehending a daily newspaper article, the clients themselves and their therapists do the rating; for adults unable to read or who are too disturbed to perform this task satisfactorily themselves, therapists and significant others perform the ratings.

After clinicians become familiar with the scales, it takes them between one and two minutes, at the end of the diagnostic interview, to fill them out. Most clients and significant others fill out the scale in five to eight minutes.

During the last eight years, numerous studies were conducted utilizing the PES scales with various client samples in order to answer questions relative to ease of use, clarity of language, and ordering and spacing of behavioral descriptions for each dimension. In addition, the psychometric properties of the scales were studied in great detail, examining such questions as score distribution, rater reliability, similarities and differences in ratings of therapist, patient, and significant other, sex and age differences in ratings, normative group ratings, and the like. The results of these studies are described in the PES manual (Ihilevich and Gleser, 1981), which can be obtained from the first author. Here we shall provide a brief outline of the evidence for generalizability, validity and usefulness of the scales.

GENERALIZABILITY STUDIES

Studies were conducted to determine the generalizability (Cronbach et al., 1972) of scores for therapists and over occasions.

AGREEMENT AMONG THERAPISTS

An early study was designed to determine to what extent therapists differ in their assessment of current status of individual clients and in setting goals for therapy. For this purpose a staff member sat in on the initial interview by the therapist and independently rated the client as to present status and three-month goals. A sample of 20 adults were rated. Data for only the first six scales are available, since the seventh scale, *Attitude Toward Self,* had not yet been fully developed.

The mean profiles for clients as rated by two staff members are reported in Table 1 and shown in Figure 1. It is evident that with the exception of one or two scales, the mean current status profiles are in close agreement. The most notable difference is in the scale for *Use of Free Time,* where the difference is .40, a statistically nonsignificant difference. The goals, on the average, are also quite close.

The data were analyzed for each scale separately to obtain estimates of the amount of variance attributable to differences among clients (σ_p^2) and that due to average differences between ratings of therapists of any one person (σ_e^2). The ratio of (σ_p^2) to the expected observed score variance (σ_x^2) yields an estimate of reliability (r_{xx}). These results are shown in Table 1. This method of computing reliability takes into account the fact that in ordinary use of these scales, clients will be rated by different therapists. Therefore, therapist mean score differences will contribute to both observed score and error variance. Such estimates, however, tend to be lower than those obtained by correlating the two sets of scores since the latter method ignores differences among therapists' means. Reliability estimates for current status ranged from .49 for *Problems* to a high of .86 for *Getting Along With Others.* The range of reliability for goals was somewhat lower (.39 to .67). The median reliabilities were .65 for present status and .43 for goals.

Since reliability estimates depend very heavily on the amount of variability among subjects in the sample, a somewhat better idea of the

TABLE 1
Comparison of Two Therapists' Ratings on an Adult Outpatient Sample[a]

Scale	*Present Status*						*Goal Rating*					
	M_{t_1}	M_{t_2}	σ_x^2	σ_p^2	σ_e^2	r_{xx}	M_{t_1}	M_{t_2}	σ_x^2	σ_p^2	σ_e^2	r_{xx}
Family Interaction	3.90	3.50	.53	.33	.20	.62	4.35	4.40	.45	.18	.28	.39
Occupation	4.30	4.35	.54	.37	.18	.68	4.75	4.75	.20	.10	.10	.49
Getting Along with Others	3.50	3.40	1.05	.90	.15	.86	4.05	4.15	.67	.45	.22	.67
Feelings and Mood	2.85	2.80	1.32	.79	.52	.60	4.05	3.95	.53	.18	.35	.34
Use of Free Time	3.40	3.00	1.68	1.28	.40	.76	4.05	3.95	1.04	.49	.55	.47
Problems	3.15	3.00	.64	.32	.32	.49	4.15	3.95	.41	.16	.25	.39

a. N = 20

M_{t_1} –mean ratings of therapist #1.

M_{t_2} –mean ratings of therapist #2.

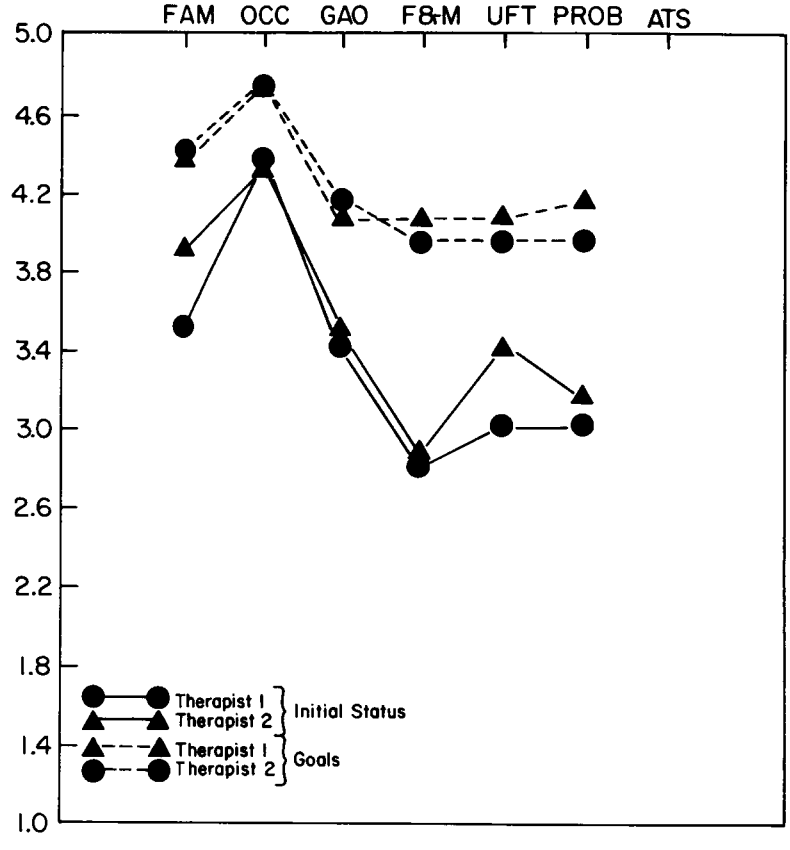

Figure 1: Comparisons of Two Therapists/Ratings for Initial Status and Goals for a Sample of Twenty Adult Outpatients

extent to which ratings of a client might differ from one therapist to another is obtained from consideration of the estimates of error variance. These range from a high of .52 to a low of .15 for ratings of current status, with a mean of .30. Furthermore, while the reliability estimates for goals were lower than for current status, the estimates of error variance were in the same range, also averaging .30. These estimates imply that 90% of the time, therapists' ratings of an individual on these

scales will lie within .9 unit from that obtainable were a large number of therapists to make every rating. Furthermore, a difference of .4 between observed means of two group of as few as 10 clients, each rated by different therapists, is reliable at the 90% level of confidence.

STABILITY OVER OCCASIONS

Stability of ratings over occasions was examined by having 65 adult patients attending group therapy sessions fill out rating scales on two occasions, two weeks apart; their therapists also rated them on both occasions.

There was little, if any, systematic variance attributable to occasions of rating on any of the scales, nor any differential trend over occasions between ratings of client and therapist (raters × occasions). Some differential trends over time were indicated for subjects (subject × occasions), particularly for *Use of Free Time, Feelings and Mood,* and *Problems,* all of which could legitimately be considered as scales measuring somewhat transient states. The triple interaction, which includes all residual random error, was also relatively large for *Feelings and Mood* and *Use of Free Time.*

The expected "true" score and error variances were computed under the assumption that ratings of either therapists or clients would be used exclusively and no generalization to other raters was intended. Error variance ranged from .16 to .37 with a mean of .22. The coefficients of generalizability over occasion for present status ranged from .54 for *Family Interaction* to .75 for *Problems* with a median of .68. For goals the range was .44 to .62 with a median of .53.

Combining information about variance among raters from the previous sample with estimates of variance over occasions from this sample, yields estimates of variance of therapists' ratings to a universe of raters and occasions of rating. These values range from .32 to .81 with a mean of .52. From these data we can deduce that 90% of the time, a therapist's ratings on the PES scales will lie within 1.2 units of that obtainable were a large number of therapists to make ratings on the same individual within a two-week interval. If higher generalizability is desirable for a particular use (e.g., individual assessment) it can be obtained by averaging ratings of two therapists, made on separate occasions; this will reduce the 90% confidence interval to ± .8.

CONSTRUCT VALIDITY STUDIES

COMPARISON OF SCORES FROM NORMATIVE AND PATIENT SAMPLES

As previously mentioned, the PES scales were constructed to evaluate clients' level of adjustment to family, job, and community, as well as personal functioning. An implication of such scaling is that well-functioning members of the community, that is, those requiring no special services, would obtain higher scores on these scales than those who are identified as emotionally disturbed. To test such an assumption, samples of male and female adults from nonpatient populations were obtained. The adult males, 90 in all, were workers in a factory (31), PTA fathers (44), and graduate students (15). The females, 171 in all, were PTA mothers (88), gainfully employed nonprofessional women (34), professional women (24), and graduate social work students (25). Additional comparison groups of physically handicapped workers were obtained from Goodwill Industries. These groups were composed of 17 males and 31 females. They ranged in age from 18 to 67 and suffered from a wide range of disabilities such as amputations, paralyses, heart disease, diabetes, and epilepsy. Educationally, they ranged from eighth grade, or less, to college graduates. These groups were compared to 111 male and 159 female clients treated on an outpatient basis at the Shiawassee County Community Mental Health Center.

The means and standard deviations of self-ratings for patient and nonpatient groups are shown in Table 2. Typically, all groups rate themselves higher on the first four scales than on the last three. These data yield important evidence for the validity of the PES scales as measures of adjustment. Self-ratings on all seven scales significantly differentiated patient and nonpatient groups; *Occupation* and *Use of Free Time* significantly differentiated the physically handicapped from nonpatient groups. All differences were in the expected direction. It is thus evident that the PES scales are capable of making valid discriminations among different groups of people with regard to their emotional, interpersonal, and community adjustment.

CORRELATIONAL STUDIES

Intercorrelations Among PES Self-Ratings. Intercorrelations among scale scores were obtained separately for the male and female normative

TABLE 2

Means and Standard Deviations of PES Self-Ratings for Patient and Non-Patient Adult Samples

Male		*Age*	*Family Interaction*	*Occupation*	*Getting Along with Others*	*Feelings and Mood*	*Use of Free Time*	*Problems*	*Attitude to Self*
Normative	X	31.2	4.65	4.84	4.73	4.69	4.31	4.36	4.38
(N = 90)	s.d.	11.7	.60	.38	.45	.53	.90	.77	.74
Physically Handicapped	X	32.5	4.18	3.94*	4.29	4.47	3.35*	4.05	4.06
(N = 17)	s.d.	12.1	1.07	1.43	.99	1.01	1.62	.97	.90
Outpatients	X	34.1	3.85**	4.01**	4.08**	3.18**	2.72**	2.73**	2.73**
(N = 111)	s.d.	12.2	.88	1.30	.92	1.28	1.36	1.20	1.14
Females									
Normative	X	31.6	4.78	4.79	4.69	4.59	3.90	4.34	4.12
(N = 171)	s.d.	9.3	.51	.52	.55	.66	.93	.73	.99
Physically Handicapped	X	38.3	4.38*	4.35*	4.48	4.29	3.26**	4.32	3.90
(N = 31)	s.d.	15.3	1.02	1.02	.62	1.04	1.26	.75	.83
Outpatients	X	31.5	4.13**	4.21**	4.14**	3.12**	2.85**	2.83**	2.48**
(N = 159)	s.d.	10.8	.71	.85	.90	1.20	1.52	1.05	1.03

*Significantly different from normative sample (Dunnett's multiple comparison procedure) $p < .05$.
**Ibid. $p < .01$.

and patient samples. While specific correlations vary somewhat from sample to sample, in general there are only low positive correlations among all the scales. The average intercorrelation for normative males is .18 and for females is .21; for outpatient males it is .24, while that for females it is .30. The fact that all correlations tend to be positive indicates that it is possible to sum scale scores to provide an overall estimate of adjustment. Their low intercorrelations indicate that the scales also yield considerable independent information.

Correlations Between Initial Status Ratings of Patient and Therapist. The correlations between the initial status scores of patient and therapist are displayed in Table 3, with therapist ratings across and client ratings down. The underlined values are the correlations between patients and therapists for the same scale; they are all highly significant ($p < .01$), averaging .48 for females and .48 for males. In both samples the lowest relationships between therapists and patients are for *Family Interaction* and *Problems;* the highest, for *Occupation, Use of Free Time,* and *Attitude Toward Self.* All off-diagonal correlations are lower than those on the diagonal. These correlations yield good evidence for the convergent and discriminant validity of the seven PES scales (Campbell and Fiske, 1959).

RELATIONSHIP TO DEMOGRAPHIC VARIABLES AND DIAGNOSTIC GROUPS

Demographic Variables. Present status and goal ratings of therapist and patient were analyzed for a group of 50 male and 50 female, sequentially drawn adult outpatients to determine differences in ratings as a function of sex, age, education, marital status, and income. Only the first six scales were available for analysis, the seventh scale not yet having been fully developed.

Only a few correlations between demographic variables and PES ratings were significant at the .05 level, and those tended to occur with ratings of males. The therapists' ratings of occupational status were higher for males who were older, better educated, in higher income brackets, and married, but these relations did not hold for ratings of females. Married men saw themselves as having better family interaction and set higher goals on this dimension than did single men; older males set lower goals for relief of problems than did younger men. For females, higher income was correlated with better family interaction.

TABLE 3

Correlations[a] Between Initial Status Ratings of Patients and Therapists for Male and Female Outpatients

	Therapist Ratings													
	Males (N = 111)							*Females (N = 159)*						
Client Ratings	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
1. Family Interaction	32	15	−07	13	12	02	08	30	22	17	12	15	07	09
2. Occupation	32	72	−16	12	13	23	12	23	62	24	24	34	12	21
3. Getting Along with Others	09	05	39	19	20	01	23	17	21	51	21	32	22	23
4. Feelings and Mood	00	20	−09	48	26	31	32	−10	23	29	40	33	23	30
5. Use of Free Time	−06	00	06	42	57	13	21	05	32	26	23	63	25	26
6. Problems	08	16	−15	24	26	38	07	−01	18	22	35	22	36	33
7. Attitude to Self	02	05	−02	33	26	10	53	06	28	25	26	23	24	53

a. Decimal points omitted.

Our studies to data indicate that demographic variables such as age, sex, marital status, income and education, have very little, if any, consistent effects on status ratings of therapists and patients on PES scales with the possible exception of *Family Interaction* and *Occupation.*

Comparison Among Diagnostic Groups. Comparisons were made among five diagnostic groups on the basis of therapists' ratings of initial status. The five diagnostic categories were: neuroses, psychoses, personality disorder, adjustment reaction, and organic brain syndrome. For each category, 25 successive patients with appropriate diagnosis were drawn, with the exception of the organic brain syndrome group which was comprised of 15 patients. The means and standard deviations of ratings of opening status for these groups are displayed in Table 4.

Multivariate analyses of variances were carried out comparing neurotics with psychotics and personality disorders with adjustment reactions. Both analyses revealed highly significant overall differences ($p < .00001$ and $p < .001$, respectively), indicating that the corresponding profiles are essentially different.

Five of the six scales differentiated neurotics and psychotics. Neurotics were rated significantly higher than psychotics on *Family Interaction, Occupation, Getting Along With Others,* and *Problems.* They were significantly lower on *Feelings and Mood.* However, only *Getting Along With Others* significantly differentiated adjustment reactions from personality disorders.

Inpatients versus Outpatients. One additional comparison that has been made is that between 25 adult inpatients from a unit in Ypsilanti State Hospital and 25 outpatients from Shiawassee Mental Health Center. For this comparison, both present status and goals were examined using therapists' ratings. The results are illustrated graphically in Figure 2.

The two groups differed significantly in present functioning by a multivariate F text. Two scales yielded significant differences—*Getting Along With Others* and *Feelings and Mood*—with inpatients rated lower. However, the most striking difference found was for goals. As clearly seen in Figure 2, the goals for each scale were set very significantly lower for inpatients than for outpatients.

SENSITIVITY TO THERAPEUTIC INTERVENTION

A number of outcome studies were conducted to ascertain the sensitivity of the scales to changes observed clinically in clients' functioning.

TABLE 4
Profiles of Initial Status of Various Diagnostic Groups[a]

	Therapist Ratings											
	Family Interaction		*Occupation*		*Getting Along with Others*		*Feelings and Mood*		*Use of Free Time*		*Problems*	
	X	*s.d.*	*X*	*s.d.*	*X*	*s.d.*	*X*	*s.d.*	*X*	*s.d.*	*X*	*s.d.*
Organic brain syndrome	2.67	1.23	2.80	1.21	2.80	.56	2.33	1.05	2.47	1.25	2.33	.82
Psychoses	3.24	.92	3.08	1.32	3.42	1.05	3.20	1.22	2.60	1.12	2.44	.77
Personality disorder	3.68	.85	3.92	1.00	3.28	.94	2.84	1.07	2.56	1.26	3.04	.73
Neuroses	4.24	.78	4.00	.82	3.76	.72	2.48	.87	2.80	1.12	3.12	.60
Adjustment reaction	4.00	.58	4.04	1.10	4.16	.69	3.16	.94	2.96	1.17	3.16	.80

a. All diagnostic groups consisted of 25 patients with the exception of Organic Brain Syndrome for which N = 15.

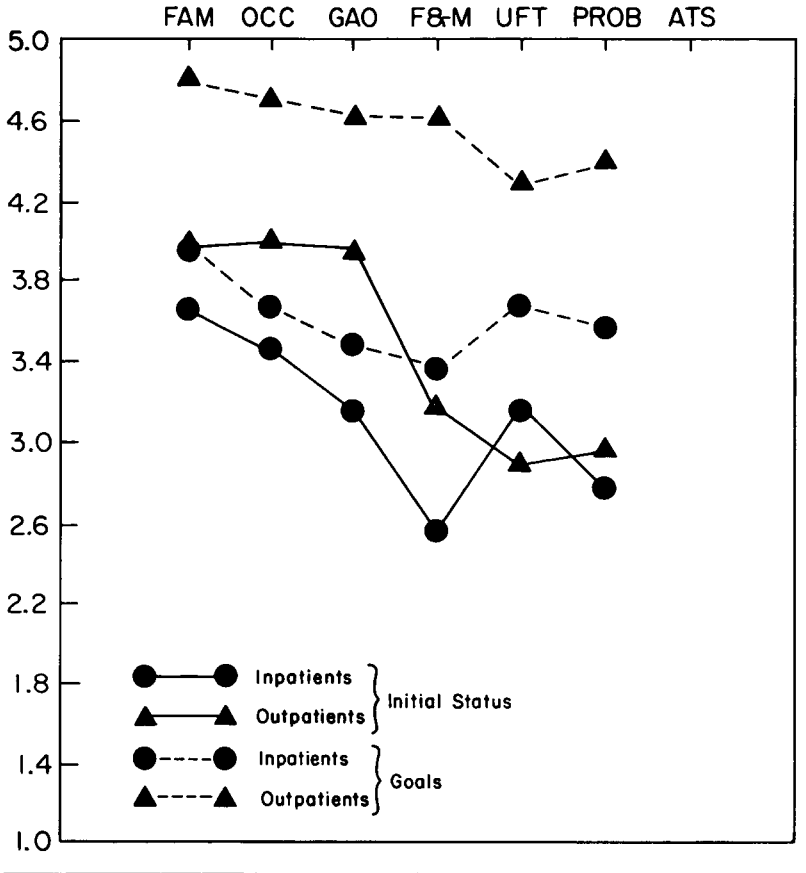

Figure 2: Comparison of Inpatients to Outpatients on Initial Status and Goals (Ratings by Therapists)

Two of these studies will be briefly described in this section. A more extensive discussion of outcome studies can be found in the manual. The first of the studies to be described here will present results found at termination of therapy; the second study will describe change in status three months after treatment was initiated for a group of clients still in psychotherapy.

The first study was carried out on a sample of 44 males and 58 females for whom initial and closing ratings were available from both patients

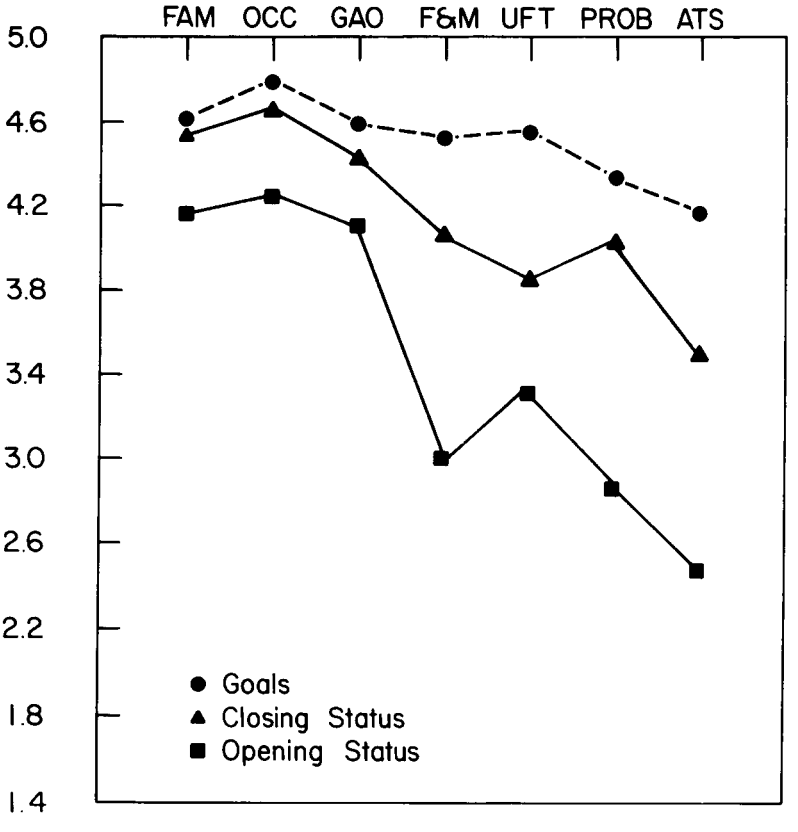

Figure 3: Initial Status, Goals, and Closing Status for 58 Female Outpatients (Ratings by Self)

and therapists. Comparing therapist ratings of closing status to opening status indicated that in their opinion both male and female patients made substantial gains in adjustment. The largest gains were obtained uniformly on *Feelings and Mood,* followed by *Problems* and *Attitude Toward Self.* It should be noted, however, that these are the areas where there is most room for improvement as judged by the initial status profiles. While three-month goals were reached at termination for only one or two scales on the average, results on the whole indicated significant improvement according to both patients and therapists on four out

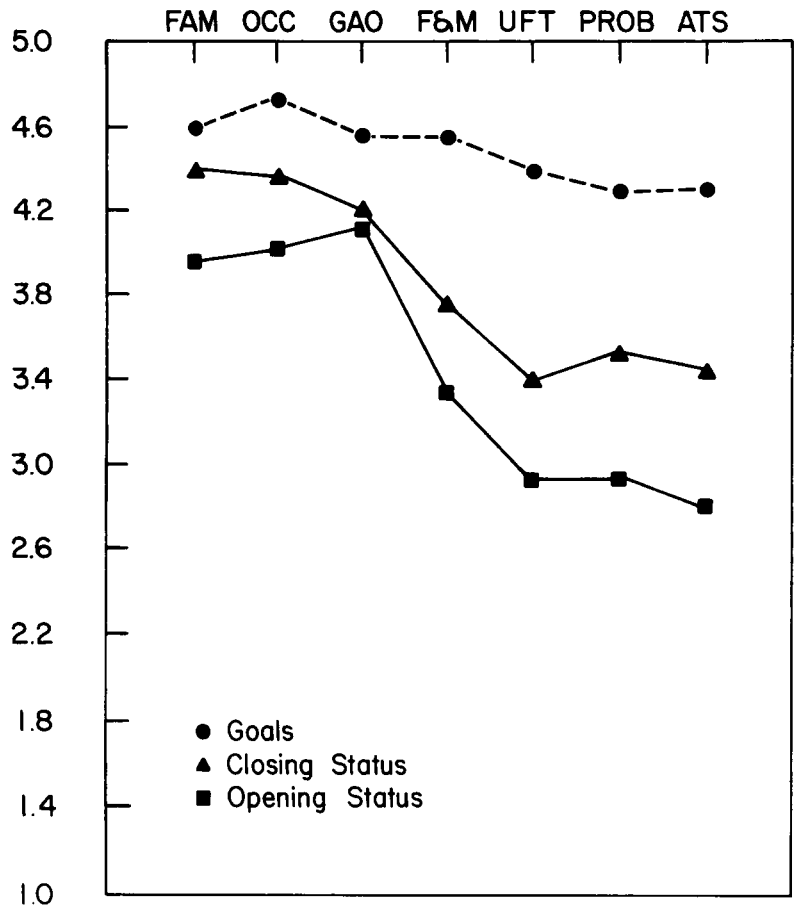

Figure 4: Initial Status, Goals, and Closing Status for 44 Male Outpatients (Ratings by Self)

of the seven scales *(Feelings and Mood, Use of Free Time, Problems,* and *Attitude Toward Self).* When the closing status ratings made by men and women clients were compared to their opening status and goals, the figures suggest that the women considered themselves somewhat more improved than did the men, particularly on *Feelings and Mood* and *Problems* (see Figures 3 and 4).

The second study involved 42 adults for whom three-month goals were set and who were then reevaluated after three months in treatment by both therapist and patient. Examining the mean status and goal ratings made by patient and therapist, initially and at three months, reveals that for these patients the most crucial problem areas are indicated by the last four scales. Their *Attitude Toward Self* was primarily negative; they complained of almost continuous *Problems;* they felt nervous, depressed or angry for days at a time *(Feeling and Mood),* and made poor *Use of Free Time.* On three of these four scales, significant improvement was made by the end of three months. For the other scales, less improvement was noted, but judging by the goals, less change was expected in three months.

THE HEURISTIC VALUE OF THE PES

The PES scales are potentially useful for exploring a broad range of issues in the clinical, programmatic, and administrative/policy domains.

In the clinical area, the system provides the possibility of obtaining continual updated information about the population being served in a manner which is economical, expeditious, and relevant to clinicians. Timely feedback can be particularly useful to therapists for review and planning functions, inasmuch as these become more and more an integral part of every mental health service delivery system. The PES scales can also be employed to investigate a broad array of applied research questions such as the impact of independent vs. mutual goal-setting by client and therapist, the impact of frequency of treatment sessions for various clincial populations, differential effectiveness of various treatment modalities for a particular clinical population, and the like. One could ask such questions as, Do clients who score low at intake on *Getting Along With Others* benefit more from attending group rather than individual therapy sessions?

In the programmatic domain, PES data should enable adminstrators and agency directors to explore questions concerned with efficient utilization of resources, selection of methodologies for optimal program outcome, cost/benefit analysis of various intervention strategies, and the like. For example, one may ask, What is the differential impact of a 12-, 18- and a 24-hour per week day treatment program on the adjustment level of low functioning, chronic schizophrenics released from state hospitals?; What are the minimal resources required to achieve certain standards of functioning in the community?

In the policy domain, the focus of interest is on planning, priority-setting, allocation of resources, and the like. Here, such questions are raised as, Is there evidence that the clients one proposes to serve benefit from the kind of program for which funds are requested?; Is there a fiscal and social advantage for the community to implement a particular program?; Does a proposed philosophy of service-delivery enhance family cohesiveness, foster its independence and self-respect?

While PES data alone cannot answer such all-encompassing questions, we propose nevertheless that this evaluation system can generate valuable information relevant to these issues which can then be weighed with other considerations in the decision-making process.

CONCLUDING REMARKS

Results of extensive studies over an eight-year period indicate that the PES is a useful measuring device for assessment of current status and change over time in personal, social, and community adjustment. The chief attributes of the scales are as follows:

(1) Interviewers can complete them in one to two minutes following a routine diagnostic interview. This factor is of considerable importance in gaining the cooperation of mental health professionals both for adopting the scales for use on a continuous basis and for maintaining high quality of gathered data.

(2) New staff from various professional disciplines joining a mental health agency can easily learn how to use the scales and integrate them into their routine work with clients.

(3) The one-page format, the colloquial language of the scales, and the simple procedures of administration make it possible for most community mental health clients, and their significant others, to fill out the scales within five to eight minutes.

(4) Comparisons of clients' ratings to those of their therapists and significant others, as well as comparisons of clients' own ratings over time, yield important ongoing information on the feelings, attitudes, and expectations of the principal people who affect the outcome of the therapeutic endeavor.

(5) Rater agreement and stability have been shown to be adequate for these brief scales, particularly for use in group comparisons. The ratings of two separate interviewers can be averaged to achieve improved reliabilty for individual comparisons.

(6) Construct validity studies reveal that the scales (a) differentiate between normal and patient groups; (b) differentiate among groups of various degrees of psychopathology; (c) are by and large independent of demographic variables; (d) meet criteria for convergent and discriminant validity on the basis of correlations between independent ratings of client and therapist; (e) measure different domains of behavior and experience as indicated by the low intercorrelations among them; and (f) are sensitive to changes in level of personal, social and community adjustment as indicated by the independent ratings of therapists and patients at the beginning, reevaluation, and termination of therapy.

(7) Finally, the PES scales have extensive heuristic value for exploring important questions in the clinical, programmatic, and policy domains.

Screening _____
Initial _____
Reeval. No. _____
Closing _____

Client _____ CR TD DX UA OT
M/F/O–Sig. Other _____ CR TD DX UA OT
Therapist _____

APPENDIX A
PROGRESS EVALUATION SCALES
(ADULT FORM)

INSTRUCTIONS – 1

Please circle one statement in each column that describes best how you were in the last two weeks.

Date _____

Name _____ Case # _____

FAMILY INTERACTION	OCCUPATION (SCHOOL, JOB OR HOMEMAKING	GETTING ALONG WITH OTHERS	FEELINGS AND MOOD	USE OF FREE TIME	PROBLEMS	ATTITUDE TOWARD SELF
Often must have help with basic needs (e.g., feeding, dressing, toilet).	Does not hold job, or care for home, or go to school.	Always fighting or destructive; or always alone.	Almost always feels nervous, or depressed, or angry and bitter, or no emotions at all.	Almost no recreational activities or hobbies.	Severe problems most of the time.	Negative attitude most of the time
Takes care of own basic needs but must have help with everyday plans and activities.	Seldom holds job, or attends classes, or cares for home.	Seldom able to get along with others without quarreling or being destructive; or is often alone.	Often feels nervous, or depressed, or angry and bitter, or hardly shows any emotion for weeks at a time.	Only occasional recreational activities, or repeats the same activity over and over again.	Severe problems some of the time or moderate problems continuously.	Negative attitude much of the time.

(continued)

APPENDIX A Continued

FAMILY INTERACTION	OCCUPATION (SCHOOL, JOB OR HOMEMAKING	GETTING ALONG WITH OTHERS	FEELINGS AND MOOD	USE OF FREE TIME	PROBLEMS	ATTITUDE TOWARD SELF
Makes own plans but without considering the needs of other family members.	Sometime holds job, or attends some classes, or does limited housework.	Sometimes quarreling, but seldom destructive; difficulties in making friends.	Frequently in a good mood but occasionally feels nervous, or depressed, or angry for days at a time.	Participates in some recreational activities or hobbies.	Moderate problems most of the time, or mild problems almost continuously.	Almost equal in positive and negative attitude toward self.
Tries to consider everyone's needs but somehow decisions and actions do not work well for everybody in the family.	Holds regular job, or classes, or does housework (or some combination of these), but with difficulty.	Gets along with others most of the time; has occasional friends.	Usually in a good mood, but occasionally feels nervous, or unhappy, or angry all day.	Often participates in recreational activities and hobbies.	Occasional moderate problems.	Positive attitude toward self much of the time.
Usually plans and acts so that own needs as well as needs of others in the family are considered.	Holds regular job, or attends classes, or does housework (or some combination of these) with little or no difficulty.	Gets along with others most of the time; has regular close friends.	In a good mood most of the time, and usually able to be as happy, or sad, or angry as the situation calls for.	Participates in, as well as creates, variety of own recreational activities and hobbies for self and others.	Occasional mild problems.	Positive attitude toward self most of the time.

Comments: _____

Rev. 4/78 SC-1104

APPENDIX B

Instructions for Goal Setting

The PES form for Goal Setting is identical to the Current Functioning form appearing in this article. The only change is in the instructions, appearing in the upper left-hand corner box of the PES form. The instructions for Goal Setting read as follows:

> INSTRUCTIONS – 2
>
> Please circle *one* statement in each column that describes best how you expect to be in _____ months.

Instructions for Handling Incomplete Ratings

When a patient or significant-other do not complete the PES form, the mental health worker indicates the reason for this omission on the body of the incompleted PES form, by circling one of the following codes, which are printed at the upper right-hand corner of the PES form:

Example:	Client:	CR	(TD)	DX	UA	OT
	Significant-Other:	CR	TD	DX	(UA)	OT

CR – cannot read
TD – too disturbed (e.g., hallucinates)
DX – seen for diagnostic purposes only (e.g., evaluation for court)
UA – unavailable (usually occurs at reevaluation or termination)
OT – other reasons (e.g., refuses)

REFERENCES

CAMPBELL, D. T. and D. W. FISKE (1959) "Convergent and discriminant validation by the multi-trait multi-method matrix." Psych. Bull. 56: 81-105.

CIARLO, J. A. and J. REIHMAN (1974) The Denver Community Mental Health Questionnaire. (Unpublished manuscript available from J. A. Ciarlo, Mental Health System Evaluation Project, 70 W. 6th Avenue, Denver, Colorado 80204).

CRONBACH, L. J., G. C. GLESER, H. NANDA and N. RAJARATNAM (1972) The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. New York: John Wiley.

FENICHEL, O. (1945) The Psychoanalytic Theory of Neurosis. New York: Norton.

FREUD, S. (1912) "Recommendations for physicians on the psychoanalytic method of treatment," in S. Freud (1959) Collected Papers, Volume 2. New York: Basic Books.

——— (1909) "Analysis of phobia in a five-year-old boy," in S. Freud (1959) Collected Papers, Volume 3. New York: Basic Books.

IHILEVICH, D. and G. C. GLESER (1981) Measuring Program Outcome: The Progress Evaluation Scales. Owosso, Michigan. (Available from D. Ihilevich, Shiawassee County Community Mental Health Services Board, 826 W. King, P.O. Box 479, Owosso, Michigan 48867)

JAHODA, M. (1958) Current Concepts of Positive Mental Health. New York: Basic Books.

KIRESUK, T. J. and R. E. SHERMAN (1968) "Goal attainment scaling: A general method for evaluating comprehensive community mental health programs." Community Mental Health Journal 4, 6: 443-453.

OFFER, D. and M. SABSHIN (1966) Normality: Theoretical and Clinical Concepts of Mental Health. New York: Basic Books.

PERLOFF, R., E. PERLOFF and E. SUSSANA (1976) "Program evaluation," in M. R. Rosenzweig and L. W. Porter (eds.) Annual Review of Psychology. Palo Alto: Annual Reviews.

SODDY, K. and R. H. AHRENFELDT [eds.] (1967) Mental Health and Contemporary Thought. London: Tavistock.

David Ihilevich received his Ph.D. in Clinical Psychology in 1968 from the University of Cincinnati. Since 1969, he has been the Director of Shiawassee County Community Mental Health Center. His main research interests are in the area of defense mechanisms and evaluation of outcome of community mental health services.

Goldine C. Gleser received her Ph.D. in Psychology from Washington University, St. Louis, in 1950. She is currently Professor Emeritus of Psychology in the Department of Psychiatry, University of Cincinnati College of Medicine. Her publications include books and articles on psychometric theory and applications, verbal behavior analysis, clinical evaluation, and disaster research.

Gordon W. Gritter received his M.D. at Boston University in 1955, and training in psychiatry at Langley Porter Institute. After several years of private practice and teaching in San Francisco, he joined the staff of Shiawassee County Community Mental Health Center, then spent several years as Consultant to the Wellington Hospital Board in New Zealand. He is now Associate Professor of Psychiatry at Michigan State University and Clinical Director of Michigan Institute for Mental Health.

Lenore J. Kroman received her M.S. from Case Western University in 1944. She is currently Associate Professor of Social Work at the University of Michigan, School of Social Work. Her clinical background and interests include inpatient, aftercare, and Community Mental Health services.

Andrew S. Watson received his M.D. in 1950 from Temple University in Philadelphia. He is Professor of Psychiatry and Professor of Law at the University of Michigan. His main interest is social psychiatry as it relates especially to the legal profession and legal institutions. He is also trained in psychoanalysis.