An Inventory for Measuring Depression

A. T. BECK, M.D.
C. H. WARD, M.D.
M. MENDELSON, M.D.
J. MOCK, M.D.
AND
J. ERBAUGH, M.D.
PHILADELPHIA

The difficulties inherent in obtaining consistent and adequate diagnoses for the purposes of research and therapy have been pointed out by a number of authors. Pasamanick[12] in a recent article viewed the low interclinician agreement on diagnosis as an indictment of the present state of psychiatry and called for "the development of objective, measurable and verifiable criteria of classification based not on personal or parochial considerations, but on behavioral and other objectively measurable manifestations."

Attempts by other investigators to subject clinical observations and judgments to objective measurement have resulted in a wide variety of psychiatric rating scales.[4,15] These have been well summarized in a review article by Lorr[11] on "Rating Scales and Check Lists for the Evaluation of Psychopathology." In the area of psychological testing, a variety of paper-and-pencil tests have been devised for the purpose of measuring specific personality traits; for example, the Depression-Elation Test, devised by Jasper[9] in 1930.

This report describes the development of an instrument designed to measure the behavioral manifestations of depression. In the planning of the research design of a project aimed at testing certain psychoanalytic formulations of depression, the necessity for establishing an appropriate system for identifying depression was recognized. Because of the reports on the low degree of interclinician agreement on diagnosis,[13] we could not depend on the clinical diagnosis, but had to formulate a method of defining depression that would be reliable and valid.

The available instruments were not considered adequate for our purposes. The Minnesota Multiphasic Personality Inventory, for example, was not specifically designed

Submitted for publication Nov. 29, 1960.

This investigation was supported by Research Grant M3358 from the National Institute of Mental Health, U.S. Public Health Service.

From the Department of Psychiatry, University of Pennsylvania School of Medicine and the Philadelphia General Hospital.

for the measurement of depression; its scales are based on the old psychiatric nomenclature; and factor analytic studies reveal that the Depression Scale contains a number of heterogeneous factors only one of which is consistent with the clinical concept of depression.[3] Jasper's Depression-Elation test[9] was derived from a study of normal college students, and his report does not refer to any studies with a psychiatric population.

Method

A. *Construction of the Inventory.*—The items in this inventory were primarily clinically derived. In the course of the psychoanalytic psychotherapy of depressed patients, the senior author made systematic observations and records of the characteristic attitudes and symptoms of depressed patients. He selected a group of these attitudes and symptoms that appeared to be specific for these depressed patients and which were consistent with the descriptions of depression contained in the psychiatric literature.[10] On the basis of this procedure, he constructed an inventory composed of 21 categories of symptoms and attitudes. Each category describes a specific behavioral manifestation of depression and consists of a graded series of 4 to 5 self-evaluative statements. The statements are ranked to reflect the range of severity of the symptom from neutral to maximal severity. Numerical values from 0-3 are assigned each statement to indicate the degree of severity. In many categories, 2 alternative statements are presented at a given level and are assigned the same weight; these equivalent statements are labeled a and b (for example, 2a, 2b) to indicate that they are at the same level. The items were chosen on the basis of their relationship to the overt behavioral manifestations of depression and do not reflect any theory regarding the etiology or the underlying psychological processes in depression.

The symptom-attitude categories are as follows:

a. Mood
b. Pessimism
c. Sense of Failure
d. Lack of Satisfaction
e. Guilty Feeling
f. Sense of Punishment
g. Self-Hate
h. Self Accusations
i. Self Punitive Wishes
j. Crying Spells
k. Irritability
l. Social Withdrawal
m. Indecisiveness
n. Body Image
o. Work Inhibition
p. Sleep Disturbance
q. Fatigability
r. Loss of Appetite
s. Weight Loss
t. Somatic Preoccupation
u. Loss of Libido

B. *Administration of the Inventory.*—The inventory was administered by a trained interviewer (a clinical psychologist or a sociologist) who read aloud each statement in the category and asked the patient to select the statement that seemed to fit him the best at the present time. In order that the instrument reflect the current status of the patient, the items were presented in such a way as to elicit the patient's attitude at the time of the interview. The patient also had a copy of the inventory so that he could read each statement to himself as the interviewer read each statement aloud. On the basis of the patient's response, the interviewer circled the number adjacent to the appropriate statement.

In addition to administering the Depression Inventory, the interviewer collected relevant background data, administered a short intelligence test, and elicited dreams and other ideational material relevant to the psychoanalytic hypotheses being investigated. These additional procedures were all administered after the Depression Inventory.

C. *Description of Patient Population.*—The patients were drawn from the routine admissions to the psychiatric outpatient department of a university hospital (Hospital of the University of Pennsylvania) and to the psychiatric outpatient department and psychiatric inpatient service of a metropolitan hospital (Philadelphia General Hospital). The outpatients were seen either on the same day of their first visit to the outpatient department or a specific appointment was made for them to come back a few days later for the complete work-up. Hospitalized patients were all seen the day following their admission to the hospital, i.e., during their first full day in the hospital. The demographic features of the population are listed in Table 1. It will be noted that there are 2 patient samples, one the original group (226 patients), the other the replication group (183 patients). The original sample (Study I) was taken over a 7-month period starting in June, 1959, the second (Study II) over a 5-month period. The completion of the first study coincided with the introduction of some new projective tests not relevant to this report.

The most salient aspects of this table are the predominance of white patients over Negro patients, the age concentration between 15 and 44, and the high frequency of patients in the lower socioeconomic groups (IV and V). The social position was derived from Hollingshead's Two Factor Index of Social Position,[7] which uses the factors of education and occupation in the class level determination.

The distribution of diagnoses was similar for Studies I and II. Patients with organic brain damage and mental deficiency were automatically excluded from the study. The proportions among the major diagnostic categories were psychotic

TABLE 1.—*Percentage Distribution of Demographic Characteristics of Patient Population*

	N	Sex		Race		Age					Social Index				Inpatient	Outpatient
		Male	Female	White	Negro	15-24	25-34	35-44	45-54	55+	I-III	IV	V			
Study I	226	40.7	59.3	67.6	32.4	24.5	33.3	25.8	11.1	5.3	13.7	40.6	45.7	33.6	66.4
Study II	183	37.2	62.8	61.2	38.8	21.3	31.2	23.5	16.9	7.1	17.4	34.8	47.8	34.4	65.6
Combined	409	39.1	60.9	64.7	35.3	23.0	32.4	24.8	13.7	6.1	15.3	38.1	46.6	34.0	66.0

disorder 41%, psychoneurotic disorder 43%, personality disorder 16%. The distributions among the subgroups were in order of frequency as follows:

	Per Cent
Schizophrenic reaction	28.2
Psychoneurotic depressive reaction	25.3
Anxiety reaction	15.5
Involutional reaction	5.5
Psychotic depressive reaction	4.7
Personality trait disturbance	4.5
Sociopathic personality	4.5
Psychophysiological disorder	3.4
Manic-depressive, depressed	1.8
Personality pattern disturbance	1.8
All other diagnoses	4.8
	100.0

D. *External Criterion.*—The patient was seen either directly before or after the administration of the Depression Inventory by an experienced psychiatrist who interviewed him and rated him on a 4-point scale for the Depth of Depression. The psychiatrist also rendered a psychiatric diagnosis and filled out a comprehensive form designed for the study. In approximately half the cases, the psychiatrist saw the patient first; in the remainder, the Depression Inventory was administered first.

Four experienced psychiatrists participated in the diagnostic study.* They may be characterized as a group as follows: approximately 12 years experience in psychiatry, holding responsible teaching and training positions, certified by the American Board of Psychiatry, interested in research, and analytically oriented.

The psychiatrists had several preliminary meetings during which they reached a consensus regarding the criteria for each of the nosological entities and focused special attention on the various types of depression. In every case, the *Diagnostic and Statistical Manual of Mental Disorders* of the American Psychiatric Association[1] was used, but it was found that considerable amplification of the diagnostic descriptions was necessary. After they had reached complete agreement on the criteria to be used in making their clinical judgments, the psychiatrists composed a detailed instruction manual to serve as a guide in their diagnostic evaluations.

The psychiatrists then participated in a series of interviews, during which two of them jointly interviewed a patient while the other two observed through a one-way screen. This served as

*In the initial group of 226 patients, some of the diagnostic evaluations were made by a "nonstandard diagnostician," that is, a psychiatrist other than the 4 regular psychiatrists. In all, 40 patients were seen by these psychiatrists.

a practical testing ground for the application of the agreed-on instructions and principles and allowed further discussion of interview techniques, the logic of diagnosis, and the pinpointing of specific diagnoses.

Since the main focus of the research was to be on depression, the diagnosticians also established specific indices to be used in making a clinical estimation of the Depth of Depression. These indices represented the pooled experience of the 4 clinicians and were arrived at independently of the Depression Inventory. For each of the specified signs and symptoms the psychiatrists made a rating on a 4-point scale of none, mild, moderate, and severe. The purpose of specifying these indices was to facilitate uniformity among the psychiatrists; however, in making the over-all rating of the Depth of Depression, they made a global judgment and were not bound by the ratings in each index.† They also concentrated on the intensity of depression at the time of the interview; hence, the past history was not as important as the mental status examination.

The indices of depression which were devised and used by the psychiatrists were as follows:

I. Apperance
- Facies
- Gait
- Posture
- Crying
- Speech
  - Volume
  - Key
  - Speed
  - Amount

II. Thought Content
- Reported Mood
- Helplessness
- Pessimism
- Feelings of Inadequacy and Inferiority
- Somatic preoccupation
- Conscious guilt
- Suicidal content

III. Vegetative Signs
- Sleep
- Appetite
- Constipation

IV. Psychosocial Performance
- Indecisiveness
- Loss of drive
- Loss of interest
- Fatigability

The diagnosticians also rated the patient on the degree of agitation and overt anxiety and filled out a check list to indicate the presence of other specific psychiatric and psychosomatic symptoms and disturbances in concentration, memory, recall, judgment, and reality testing. He also made a rating of the severity of the present illness on a 4-point scale.

†A number of problems arose in assessing the relative degree of depression of patients with contrasting clinical pictures. For example, would a patient who is regressed and will not eat be rated as more depressed than a patient who is not regressed but has made a genuine suicidal attempt? Such problems involved complex clinical judgments and will be the subject of a later report.

To establish the degree of agreement, the psychiatrists interviewed 100 patients and made independent judgments of the diagnosis and the Depth of Depression. All 4 diagnosticians participated in the double assessment and were randomly paired with one another so that each of the patients was seen by 2 diagnosticians. The procedure was to have one psychiatrist interview the patient and then after a resting period of a few minutes, the other psychiatrist would interview the patient. After the second interview was complete, the clinicians generally would meet and discuss the cases seen concurrently to ascertain the reasons for disagreement (if any).

Results

A. *Reliability of Psychiatrists' Ratings.*—The agreement among the psychiatrists regarding the major diagnostic categories of psychotic disorder, psychoneurotic disorder, and personality disorder was 73% in the 100 cases that were seen by 2 psychiatrists.‡ This level of agreement, while higher than that reported in many investigations, was considered too low for the purposes of our study.

The degree of agreement, however, in the rating of "Depth of Depression" was much higher. Using the 4-point scale (none, mild, moderate, and severe) to designate the intensity of depression, the diagnosticians showed the following degree of agreement:

Complete agreement	56%
One degree of disparity	41%
Two degrees of disparity	2%
Three degrees of disparity	1%

This indicates that there was agreement within one degree on the 4-point scale in 97% of the cases.

B. *Reliability of Depression Inventory.*—Two methods for evaluating the internal consistency of the instrument were used. First, the protocols of 200 consecutive cases were analyzed. The score for each of the 21 categories was compared with the total score on the Depression Inventory for each individual. With the use of the Kruskal-Wallis Non-Parametric Analysis of Vari-

‡A detailed description of the reliability studies will be reported in a separate article.[3] The types of disagreement regarding the nosological categories and the reasons for disagreement are being systematically investigated in another study.

ance by Ra
categories sh
to the total s
nificance was
categories ex
category), wl
level.

The secon
sistency was
half reliabili
first sample
The Pearson
categories w
liability coef
man-Brown
to 0.93.[5]

Certain tr
the stability
and questio
method and
were not a
the Depressi
reasons: If
istered afte
correlation
be spuriousl
factor. If
the consiste
of the fluct
pression th
The same
administrati
terviewers.

Two ind
stability of
The first w
method. Th
a group o
times. At
of the test.

§ This pro
variation in
associated w
inventory. I
total invent
particular
The Kruska
whether the
increased si
ences in se
alternative

ance by Ranks,[14] it was found that all categories showed a significant relationship to the total score for the inventory.§ Significance was beyond the 0.001 level for all categories except category S (Weight-loss category), which was significant at the 0.01 level.

The second evaluation of internal consistency was the determination of the split-half reliability. Ninety-seven cases in the first sample were selected for this analysis. The Pearson r between the odd and even categories was computed and yielded a reliability coefficient of 0.86; with a Spearman-Brown correction, this coefficient rose to 0.93.[5]

Certain traditional methods of assessing the stability and consistency of inventories and questionnaires, such as the test-retest method and the inter-rater reliability method, were not appropriate for the appraisal of the Depression Inventory for the following reasons: If the inventory were readministered after a short period of time, the correlation between the 2 sets of scores could be spuriously inflated because of a memory factor. If a long interval was provided, the consistency would be lowered because of the fluctuations in the intensity of depression that occur in psychiatric patients. The same factors precluded the successive administration of the test by different interviewers.

Two indirect methods of estimating the stability of the instrument were available. The first was a variation of the test-retest method. The inventory was administered to a group of 38 patients at two different times. At the time of each administration of the test, a clinical estimate of the Depth of Depression was made by one of the psychiatrists. The interval between the 2 tests varied from 2 to 6 weeks. It was found that changes in the score on the inventory tended to parallel changes in the clinical Depth of Depression, thus indicating a consistent relationship of the instrument to the patient's clinical state. (These findings are discussed more fully in the section on validation studies.)

An indirect measure of inter-rater reliability was achieved as follows: Each of the scores obtained by each of the 3 interviewers was plotted against the clinical ratings. A very high degree of consistency among the interviewers was observed for the mean scores respectively obtained at each level of depression. Curves of the distribution of the Depression Inventory scores plotted against the Depth of Depression were notably similar, again indicating a high degree of correspondence among those who administered the inventory.

C. *Validation of the Depression Inventory.*—The means and standard deviations for each of the Depth of Depression categories are presented in Table 2. It can be seen from inspection that the differences among the means are as expected; that is, with each increment in the magnitude of depression, there is a progressively higher mean score. The Kruskal-Wallis One-way Analysis of Variance by Ranks was used to evaluate the statistical significance of these differences; for both the original group (Study I) and the replication group (Study II), the p-value of these differences is <0.001.

Since the Kruskal-Wallis test evaluates the over-all association between the scores on the Depression Inventory and the Depth of Depression ratings, the Mann-Whitney U test[14] was used to appraise the power of the Depression Inventory to discriminate between specific Depth of Depression categories. It was found that all differences between adjacent categories (none, mild, moderate, and severe) in both studies were significant at <0.0004 with the exception of

§ This procedure is designed to assess whether variation in response to a particular category is associated with variation in total score on the inventory. For each category, the distribution of total inventory scores for individuals selecting a particular alternative response was determined. The Kruskal-Wallis test was then used to assess whether the ranks of the distribution of total scores increased significantly as a function of the differences in severity of depression indicated by these alternative responses.

TABLE 2.—*Distribution of Means and Standard Deviations of Depression Inventory Scores as a Function of Depth of Depression Ratings*

		Clinical Rating of Depth of Depression										
		0 (None)		1 (Mild)		2 (Moderate)		3 (Severe)				
	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N	p *
Study I	226	11.3	7.7	63	19.0	9.9	70	24.7	10.2	73	28.0	9.4	20	<0.001
Study II	183	10.3	8.6	52	18.1	10.1	57	26.1	8.9	61	32.7	11.5	13	<0.001
Combined I & II	409	10.9	8.1	115	18.7	10.2	127	25.4	9.6	134	30.0	10.6	33	

* Kruskal-Wallis One-Way Analysis of Variance.

TABLE 3.—*Correlation Between Depression Inventory Scores and Clinicians' Ratings of Depth of Depression*

	n	Correlation Coefficient *	Standard Error	p
Study I	226	0.65	0.068	<0.01
Study II	183	0.67	0.059	<0.01

* Pearson biserial r.

the differences between the moderate and severe categories, which had a p-value of <0.1 in Study I and <0.02 in Study II.

A Pearson biserial r [6] was computed to determine the degree of correlation between the scores on the Depression Inventory and the clinical judgment of Depth of Depression. To perform this correlation, the criterion ratings were reduced from 4 to 2 (none and mild, moderate and severe). The obtained biserial coefficients are highly significant as shown in Table 3.

Another index of the power of the inventory to distinguish among groups is provided by the computation of false negatives and false positives when the Depression Inventory scores are plotted against the clinical ratings of Depth of Depression. In view of the fact that the psychiatrists' ratings overlap considerably in adjacent Depth-of-Depression categories, it was decided that for this analysis the Depression Inventory scores in nonadjacent categories should be compared. The procedure employed was as follows: The data in Study I were analyzed and cutting scores were established. The same cutting scores were used for Study II. The results are shown in Table 4.

It will be noted that there are fewer false negatives and false positives in Study II than in Study I. This may be accounted for by the fact that in Study II only the 4 "standard" psychiatrists were used and by the fact that with increasing experience in assessing the severity of depression, they achieved greater precision in their judgments. As expected, the most clear-cut discriminations occurred when extreme groups (none vs. severe) were compared. In Study I the cutting score discriminated between

Between Depression
Clinicians' Ratings
Depression

n t *	Standard Error	p
	0.068	<0.01
	0.059	<0.01

n the moderate and
ch had a p-value of
<0.02 in Study II.
r [6] was computed to
f correlation between
ression Inventory and
of Depth of De-
this correlation, the
reduced from 4 to
oderate and severe).
coefficients are highly
in Table 3.
the power of the in-
h among groups is
utation of false nega-
ves when the Depres-
s are plotted against
Depth of Depression.
that the psychiatrists'
siderably in adjacent
categories, it was de-
nalysis the Depression
nonadjacent categories
. The procedure em-
s: The data in Study
l cutting scores were
e cutting scores were
The results are shown

it there are fewer false
positives in Study II
s may be accounted for
Study II only the 4
ists were used and by
ncreasing experience in
ty of depression, they
ecision in their judg-
the most clear-cut dis-
d when extreme groups
ere compared. In Study
discriminated between

TABLE 4.—*Effectiveness of Depression Inventory Cutting Scores in Discriminating Between Nonadjacent Depth of Depression Categories*

	Study I						Study II					
	Depth of Depression Categories						Depth of Depression Categories					
	None vs Mod. C.S.=17		Mild vs Sev. C.S.=26		None vs Sev. C.S.=21		None vs Mod. C.S.=17		Mild vs Sev. C.S.=2		None vs Sev. C.S.=21	
N above C.S.	14	56	16	13	5	15	8	53	11	9	4	11
N below C.S.	49	17	54	7	58	5	44	8	46	4	48	2
	—	—	—	—	—	—	—	—	—	—	—	—
Total N	63	73	70	20	63	20	52	61	57	13	52	13

* C.S.=cutting score on Depression Inventory.

those 2 categories in 73 out of 83 cases (88%) and in Study II in 59 out of 65 cases (91%).

A pertinent test of the inventory's power to assess the intensity of depression is its ability to reflect changes after a time interval. A group of 38 hospital patients who had received the complete work-up including the Depression Inventory and the clinical diagnostic evaluation on the first full day in the hospital were examined a second time by the same psychiatrist and received the same battery of tests. The time interval between the 2 tests ranged from 2 to 5 weeks. In 5 cases, the psychiatrist found that the changes were not major enough to warrant changing the patient from one Depth of Depression category to another; he was aware, however, of finer changes in the severity of depression in these cases. In 33 cases, there was enough gross change in the clinical picture to warrant a change from one Depth of Depression category to another. The Depression Inventory scores changed in all cases; this was consistent with the expectation that the Depression Inventory would reflect minor changes, since its range is much greater than the clinical rating scale.

Table 5 shows the results of the determination of the number of cases in which a change in the Depth of Depression was predicted by a change in the Depression Inventory score; in 28 out of the 33 cases (85%), the change in the clinical Depth of Depression was correctly predicted.

Comment

The Depression Inventory was subjected to a variety of tests to determine its re-

TABLE 5.—*Relationship of Changes in Depth of Depression Rating to Changes in Depression Inventory Scores*

		Depth of Depression	
		Decreased	Increased
Depression Inventory Score	Decreased	26	2
	Increased	3	2

liability and validity. The high correlation coefficient on the split-half item analysis and the significant relationship between the individual category scores and the total scores indicate the instrument is highly reliable. The highly significant relationship between the scores on the inventory and the clinical ratings of Depth of Depression and the power to reflect clinical changes in the Depth of Depression attest to the validity of this instrument.

When the question arises of assessing some diagnostically relevant behaviors as, for example, are presented by states of depression, the clinician is quite naturally disposed to rely upon clinical observation and tends to mistrust personality inventories. This objection to so-called "objective" instruments is formally expressed by Horn,[8] who, in commenting on the "relative sterility" of personality inventories in predicting behavior, challenges the assumption that the items in an inventory convey the same or similar meaning to everyone who takes the test. He argues that "a personality self-rating questionnaire is in the nature of a projective test: each item serves as an ambiguous stimulus whose interpretation is affected by the subject's needs, wishes, fears, etc." This approach, in his opinion, removes from consideration the efficacy of a self-rating inventory as an accurate self-evaluation. However, the adequacy of any test as an accurate index of what it is supposed to measure is essentially an empirical question and cannot be resolved by fiat. Thus, in the case of the Depression Inventory, it was possible to obtain self-evaluations from the patient that were consistent with the total behavior of the patient as observed by the clinician.

A formidable problem in evaluating the validity of an inventory centers around the adequacy of the external criterion. In view of the well known variability of psychiatric diagnoses, it is necessary to have some other consistent standard against which the inventory score can be judged. In our study, for example, there was concurrence on the primary diagnosis of depression only 50% of the time.[2] This problem was solved by having the diagnosticians make judgments of the intensity of depression. When this was done, a high degree of consistency was found among the psychiatrists' ratings. While this classification disregarded the primary diagnosis, it did employ the same diagnostic signs and symptoms that are generally considered characteristic of primary depression. The change from the usual diagnostic procedure was to treat depression as a personality dimension and not simply as a discrete nosological entity. It was found, moreover, that this particular cluster of symptoms occurred in association with almost every other nosological category. In fact, in only about 26% of the cases was depression found to be completely absent.

While the psychiatrists showed a close agreement on the estimate of the Depth of Depression, one can still raise questions as to whether they were actually assessing depression, or whether it might have been some other personality variable. While there is no reason to assume that clinical evaluation is the ultimate criterion, as long as one is dealing with a clinical phenomenon we will have to rely on expert judgment as our criterion until other measures are developed.

The ability of the inventory to approximate clinical judgments of intensity of depression offers a number of advantages in its use for research purposes. First, it meets the problem of the variability of clinical judgment of nosological entities and provides a standardized, consistent measure that is not sensitive to the theoretical orientation or idiosyncrasies of the individual who administers it. Second, since the inventory can be administered by an interviewer who is easily trained in its use, it is far more economical than a clinical psychiatric interview. Third, since the inventory provides a numerical score, it facilitates comparison with other quantitative data. Finally, since the inventory reflects changes in the Depth of Depression over time, it provides an objective measure for judging improvement

resulting from psychotherapy, drug therapy, and other forms of treatment.

While this instrument is aimed at registering varying degrees of depression along a continuum, it is not designed to distinguish among standard diagnostic categories. A regressed schizophrenic, for example, might receive the same score as a case of involutional psychosis (provided they has the same level of depression). A further limitation of the instrument is that its applicability depends on the cooperation of the patient as well as his ability to comprehend the items.

Summary

The present study describes an inventory which has been developed to provide a quantitative assessment of the intensity of depression. This instrument was administered by an interviewer to a random sample of 226 clinic and hospitalized psychiatric patients. For purposes of replication, a second sample of 183 cases was subjected to the same procedure. Independent clinical ratings of the Depth of Depression were made by experienced psychiatrists.

Studies of the internal consistency and stability of the instrument indicate a high degree of reliability. Comparisons between the scores on the inventory and the clinical judgments by the diagnosticians indicate a high degree of validity.

The inventory was able to discriminate effectively among groups of patients with varying degrees of depression. It also was able to reflect changes in the intensity of depression after an interval of time. In view of these attributes of reliability and validity, this instrument is presented as a useful tool for research study of depression, and as a step in the direction of placing psychiatric diagnosis on a quantitative basis.

The authors wish to express their appreciation to their consultants, Seymour Feshbach, Ph.D. and Marvin Hurvich, Ph.D.

Department of Psychiatry, Research Laboratories, Hospital of the University of Pennsylvania, 3400 Spruce St., Philadelphia 4.

REFERENCES

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Washington, D.C., American Psychiatric Association, 1952.

2. Beck, A. T.; Ward, C. H.; Mendelson, M.; Mock, J., and Erbaugh, J.: Unpublished study, 1960.

3. Comrey, A. L.: A Factor Analysis of Items on the MMPI Depression Scale, Educ. Psychol. Meas. 17:578-585 (Winter) 1957.

4. Goodrich, D. W.: Quantification of the Severity of Overt Psychotic Symptoms, Amer. J. Psychiat. 110:334-341 (Nov.) 1953.

5. Guilford, J. P.: Fundamental Statistics in Psychology and Education, New York, McGraw-Hill Book Company, Inc., 1956.

6. Guilford, J. P.: Psychometric Methods, Ed. 2, New York, McGraw-Hill Book Company, Inc., 1954.

7. Hollingshead, A. B.: Two Factor Index of Social Position (Mimeographed), New Haven, A. B. Hollingshead, 1957.

8. Horn, D.: Intra-Individual Variability in the Study of Personality, J. Clin. Psychol. 6:43-47 (Jan.) 1950.

9. Jasper, H. H.: A Measurement of Depression-Elation and its Relation to a Measure of Extraversion-Intraversion, J. Abnorm. Soc. Psychol. 25: 307-318 (Oct.-Dec.) 1930.

10. Kraines, S. N.: Mental Depressions and Their Treatment, New York, The Macmillan Company, 1957.

11. Lorr, M.: Rating Scales and Check Lists for the Evaluation of Psychopathology, Psychol. Bull. 51:119-127 (March) 1954.

12. Pasamanick, B.; Dintz, S., and Lefton, M.: Psychiatric Orientation and Its Relation to Diagnosis and Treatment in a Mental Hospital, Amer. J. Psychiat. 116:127-132 (Aug.) 1959.

13. Schmidt, H. O., and Fonda, C. P.: The Reliability of Psychiatric Diagnoses: A New Look, J. Abnorm. Soc. Psychol. 52:262-267 (March) 1956.

14. Siegel, S.: Nonparametric Statistics for the Behavioral Sciences, New York, McGraw-Hill Book Company, Inc., 1956.

15. Wittenborn, J. R.: Psychiatric Rating Scales, New York, The Psychological Corporation, 1955.

Appendix
Depression Inventory

A (Mood)

0 I do not feel sad

1 I feel blue or sad

2a I am blue or sad all the time and I can't snap out of it

2b I am so sad or unhappy that it is very painful
3 I am so sad or unhappy that I can't stand it

B (Pessimism)
0 I am not particularly pessimistic or discouraged about the future
1a I feel discouraged about the future
2a I feel I have nothing to look forward to
2b I feel that I won't ever get over my troubles
3 I feel that the future is hopeless and that things cannot improve

C (Sense of Failure)
0 I do not feel like a failure
1 I feel I have failed more than the average person
2a I feel I have accomplished very little that is worthwhile or that means anything
2b As I look back on my life all I can see is a lot of failures
3 I feel I am a complete failure as a person (parent, husband, wife)

D (Lack of Satisfaction)
0 I am not particularly dissatisfied
1a I feel bored most of the time
1b I don't enjoy things the way I used to
2 I don't get satisfaction out of anything any more
3 I am dissatisfied with everything

E (Guilty Feeling)
0 I don't feel particularly guilty
1 I feel bad or unworthy a good part of the time
2a I feel quite guilty
2a I feel bad or unworthy practically all the time now
3 I feel as though I am very bad or worthless

F (Sense of Punishment)
0 I don't feel I am being punished
1 I have a feeling that something bad may happen to me
2 I feel I am being punished or will be punished
3a I feel I deserve to be punished
3b I want to be punished

G (Self Hate)
0 I don't feel disappointed in myself
1a I am disappointed in myself
1b I don't like myself
2 I am disgusted with myself
3 I hate myself

H (Self Accusations)
0 I don't feel I am any worse than anybody else
1 I am very critical of myself for my weaknesses or mistakes
2a I blame myself for everything that goes wrong
2b I feel I have many bad faults

I (Self-punitive Wishes)
0 I don't have any thoughts of harming myself
1 I have thoughts of harming myself but I would not carry them out
2a I feel I would be better off dead
2b I have definite plans about committing suicide
2c I feel my family would be better off if I were dead
3 I would kill myself if I could

J (Crying Spells)
0 I don't cry any more than usual
1 I cry more now than I used to
2 I cry all the time now. I can't stop it
3 I used to be able to cry but now I can't cry at all even though I want to

K (Irritability)
0 I am no more irritated now than I ever am
1 I get annoyed or irritated more easily than I used to
2 I feel irritated all the time
3 I don't get irritated at all at the things that used to irritate me

L (Social Withdrawal)
0 I have not lost interest in other people
1 I am less interested in other people now than I used to be
2 I have lost most of my interest in other people and have little feeling for them
3 I have lost all my interest in other people and don't care about them at all

M (
0
1
2
3

N (I
0
1
2
3

O (W
0
1a
1b
2
3

P (Sl
0
1
2
3

M (Indecisiveness)
- 0 I make decisions about as well as ever
- 1 I am less sure of myself now and try to put off making decisions
- 2 I can't make decisions any more without help
- 3 I can't make any decisions at all any more

N (Body Image)
- 0 I don't feel I look any worse than I used to
- 1 I am worried that I am looking old or unattractive
- 2 I feel that there are permanent changes in my appearance and they make me look unattractive
- 3 I feel that I am ugly or repulsive looking

O (Work Inhibition)
- 0 I can work about as well as before
- 1a It takes extra effort to get started at doing something
- 1b I don't work as well as I used to
- 2 I have to push myself very hard to do anything
- 3 I can't do any work at all

P (Sleep Disturbance)
- 0 I can sleep as well as usual
- 1 I wake up more tired in the morning than I used to
- 2 I wake up 1-2 hours earlier than usual and find it hard to get back to sleep
- 3 I wake up early every day and can't get more than 5 hours sleep

Q (Fatigability)
- 0 I don't get any more tired than usual
- 1 I get tired more easily than I used to
- 2 I get tired from doing anything
- 3 I get too tired to do anything

R (Loss of Appetite)
- 0 My appetite is no worse than usual
- 1 My appetite is not as good as it used to be
- 2 My appetite is much worse now
- 3 I have no appetite at all any more

S (Weight Loss)
- 0 I haven't lost much weight, if any, lately
- 1 I have lost more than 5 pounds
- 2 I have lost more than 10 pounds
- 3 I have lost more than 15 pounds

T (Somatic Preoccupation)
- 0 I am no more concerned about my health than usual
- 1 I am concerned about aches and pains *or* upset stomach *or* constipation *or* other unpleasant feelings in my body
- 2 I am so concerned with how I feel or what I feel that it's hard to think of much else
- 3 I am completely absorbed in what I feel

U (Loss of Libido)
- 0 I have not noticed any recent change in my interest in sex
- 1 I am less interested in sex than I used to be
- 2 I am much less interested in sex now
- 3 I have lost interest in sex completely