WEYBURN ASSESSMENT SCALE*

By

D. B. BLEWETT, Ph.D.

Supervising Psychologist
Psychiatric Services Branch, Department of Public Health, Saskatchewan, Canada

and

W. B. STEFANIUK, B.A.

Research Psychologist Associate
Saskatchewan Committee for Schizophrenia Research

SOME time ago the Department of National Health and Welfare of the Government of Canada became interested in the therapeutic effects of a preparation of nucleotides in the treatment of mental illness. It was decided that the research group from the Saskatchewan Hospital, Weyburn, should assess the usefulness of the preparation in the treatment of chronic schizophrenia.

In this study, as in any attempt to evaluate the effect of a therapeutic procedure, one of the principal difficulties to be overcome was that of establishing a criterion against which change in the degree of psychiatric illness of the subjects could be measured.

One of the most common methods of assessing change after therapy is to determine the number of individuals who are discharged from hospital subsequent to the treatment under investigation. This method is of limited value since the criterion is too gross. Clear-cut changes may occur with regard to such variables as the ease of nursing care or the acting out of hostile impulses, but such changes cannot be assessed when discharge is the sole criterion of improvement.

A second method commonly used to measure change is the classification of such change as "worse", "much worse", "improved", "much improved", and "recovered". In his criticism of this method Malamud (10) has pointed out that it fails to indicate the extent to which the patient's present mental status differs from his pre-morbid condition, and that the degree of impairment is measured against a theoretical average, which involves us in the time-worn problem of describing "normality" when applied to human behaviour.

For the purpose of this study it was felt that the criterion measure should be in the form of a rating scale which would, as far as possible, yield objective, valid and reliable assessments of the degree of illness, that we should aim at the inclusion of all the rateable features of chronic schizophrenia. In short, we should attempt to develop a scale which would be essentially an operational definition of that condition.

A review of the literature dealing with the objective assessment of therapeutic change in hospitalized psychiatric patients failed to offer any measuring

* The authors wish to express their gratitude for the aid given to them by Doctors H. Osmond, J. Lucy and J. Smythies, and by many other members of the staff of the Saskatchewan Hospital, Weyburn.

This article has been approved by the Saskatchewan Committee for Schizophrenia Research.

instrument which seemed to meet the conditions of the setting in which the study was to be carried out. Rating scales such as those of Plant (1922), Wilcox (1942), Cohen *et al.* (1944), Scherer (1951), Lorr (1953) and Baker and Thorpe (1956) were devised chiefly for assessment of patients by the nursing staff, whereas in the present study, ratings were to be made by trained psychiatrists. While Wittman (1948) and Malamud and Sands (1947) had produced scales which were devised for use by psychiatrists, it was felt that these scales, which had been developed to provide assessments across the wide spectrum of schizophrenia, did not deal in sufficient detail with the specific area of chronic schizophrenia. Further, the methods of assessment reviewed generally involved pre-requisites which could not be met in our hospital. Despite the fact that we drew heavily on Malamud's experience, and became indebted to him for his discussion on the theory and application of rating scales, we found our facilities to be inadequate for the use of his form, which utilizes as its base line the premorbid level of personality and measures change as "worse" or "better" from this level. His method is an excellent one if two conditions can be met—(*a*) the case histories available and interviews conducted and reviewed are extremely full and reliable, and (*b*) a full and supplementary staff is available to carry out the task. In our case we found gross deficiencies in our case records, especially in the files of those patients who had been hospitalized many years. We did not have the staff to meet the requirements of the method, particularly as regards social workers. We were thus confronted with the need for an objective, reliable method of rating which would permit relatively sensitive assessment of all possible facets of chronic hospitalized schizophrenia across the range of severity of illness and which could be utilized with a minimum of staff.

The Initial Form

A tentative assessment form was prepared. This form included seven broad behavioural categories (appearance, behaviour, thought, perception, mood, memory and insight). These categories were broken down into 49 variables or items (for instance, appearance was assessed on the basis of the nine variables or items: cleanliness of dress, shoes, cleanliness of nails, cleanliness of hands and face, hair, general order of dress, shaving, wetness and smell). Each of these items was scored on a three-point scale. Since in this study we were concerned with change occurring in a specific experimental situation, we were able to meet Malamud's criticism against measuring from the theoretical average by establishing as a base line the level of behaviour exhibited by the patient immediately prior to treatment. The use of this method permits measurement in any direction of changes which take place during a period of therapy.

A meeting of all the doctors in the hospital was called and the separate scoring items were discussed. The search for general agreement proved to be interesting but arduous, for the medical staff represented many frequently divergent points of view.

The Pilot Study

The initial scale was used in a small pilot project aimed at determining its adequacy. In the pilot study, five patients were rated by each of three psychiatrists. The project was carried out in one day with as little time as possible between interviews.

Results of the Pilot Study

The results of this trial provided the material upon which the first revision was based. The scale was assessed in terms of the extent to which the raters found it to be clinically adequate and in terms of the level of inter-rater agreement. Each item and each behavioural category was examined for inter-rater reliability to permit modification or rejection of those items in which agreement proved to be low.

Changes and additions were made to clarify the definitions and lessen the likelihood of subjective variation in interpretation. Instructions were set down as a guide to the assessor, particularly during his first few attempts at using the scale. The raters felt that it was not until after some three or four assessments that they could work without the guide. The problem of prompting was a very important one. Where prompting was considered permissible, the number of times a subject might be prompted for each scoring item was specified and the specific method of prompting was set down. Each scoring point was defined by an item description in order to minimize subjectivity, and, to permit finer differentiation, the system was extended to a five-point scale.

W.A.S. First Revision

The revised scale contained the same seven behavioural categories as did the initial form. In this revision eleven items were dropped from the scale. For eight of these, no agreement as to interpretation could be reached and scoring differed markedly between raters. Three further items were found to duplicate material covered elsewhere in the scale.

The average time for the completion of the assessment was approximately twelve minutes. Various scoring items were re-arranged such that they follow the sequence of a "model" interview, which facilitated scoring and shortened interview time.

The Nucleotide Study

It was in this revised form that the scale was used in the nucleotide study, which was in fact its *raison d'être*. In this project the revised scale was administered independently on three occasions by each of three psychiatrists to a group of forty male chronic schizophrenic patients, all of whom had been hospitalized for over five years.

Aside from the assessment of the therapeutic effects of the nucleotide preparation, which has been reported elsewhere (9) the data were further examined with two objects in view.

1. We sought to assess statistically the reliability and validity of the rating scale and to determine, if possible, those general areas or behavioural categories and those specific items in which the application of a rating method seemed appropriate. In other words, we hoped that the statistical analysis would provide us with data basic to further revision and improvement of the rating scale.
2. We were seeking to obtain further understanding of the structure of chronic hospitalized schizophrenia through examination of the factorial structure underlying the ratings, which constitute an operational definition of that condition.

Analysis of the Data from the Nucleotide Study

It is usual to assess the validity of a scale designed to measure degrees of a given illness by initially demonstrating its ability to differentiate between the "target" syndrome group and other populations. The nature of this scale limits its use in the present form almost exclusively to intensive measurement of variation within the chronic schizophrenia category and largely precludes any extensive study involving comparisons with other conditions.

This fact appears to pose a problem in terms of validation, since we cannot demonstrate that we are measuring chronic schizophrenia but must make that assumption. However, we believe this assumption to be justified since the criterion against which validity must ultimately be assessed in studies of psychopathological conditions is psychiatric judgment and diagnosis and it is specifically those behavioural phenomena upon which the diagnosis of chronic schizophrenia is established that have been incorporated in this assessment scale.

Rather, the problem becomes one of internal validation—the determination of whether all the behavioural categories sampled are measures of the same process (e.g., whether they inter-correlate positively). This phase of the analysis was carried out in Ottawa by Dr. J. Donnelly (1953) of the Dominion Bureau of Statistics. The results are summarized in Table I.

Table I

Inter-correlation Coefficients of Behaviour Category "Factor" Scores

	Hallucination	Behaviour	Thought	Memory	Appearance	Insight	Mood
Hallucination	—	−·170	−·014	−·04	·002	−·022	·040
Behaviour ..	−·170	—	·868	·850	·688	·684	·335
Thought ..	−·014	·868	—	·834	·667	·650	·301
Memory ..	−·040	·850	·834	—	·684	·658	·258
Appearance	·002	·688	·667	·684	—	·336	·364
Insight ..	−·022	·684	·650	·658	·336	—	·312
Mood ..	·040	·335	·301	·258	·364	·312	—

The remainder of the statistical analysis was carried out by the authors.

The nature of the data is such that the principal concern must be with the question of reliability. The data lend themselves readily to such evaluation and the reliability of the obtained ratings has been assessed in the following way:

1. Inter-rater coefficients were computed for each of the three ratings to determine how well raters agreed with each other on different occasions. These results are summarized in Table II.

Table II

Inter-rater Consistency Coefficients for Three Ratings by Items

Rater	X				Y			
	1st	2nd	3rd	Total	1st	2nd	3rd	Total
Y	·70	·76	·80	·73	—	—	—	—
Z	·73	·71	·72	·72	·74	·77	·79	·77

2. Inter-rater coefficients were computed for each of the behavioural categories to assess how well the raters agreed in their ratings of each type of behaviour. These results are summarized in Table III.

TABLE III

Inter-rater Consistency Coefficients for Behaviour Scores on First Rating

	X		Y
	Z	Y	Z
Appearance	·733	·764	·771
Behaviour	·890	·801	·910
Mood	·047	−·206	·247
Thought	·768	·724	·801
Memory	·969	·953	·985
Hallucinations	·351	·333	·260
Insight	·711	·610	·599

3. An item analysis was carried out in which the scores for each item subsumed in a behavioural category were correlated with the total scores for that category. These results are summarized in Table IV.

TABLE IV

Behaviour		Appearance		Thought		Memory		Mood	
Posture	·84	Nails	·71	Sentence Structure	·90	Ward	·91	Appropriate-ness	·69
Rapport	·79	Clean (hands and face)	·70	Associative Process	·88	Birth and Age	·91	Dominant	·56
Obedience	·79	Shoes	·65	Thought Blocking	·88	Hospital	·87	Mood Tempo	·32
Gait	·77	Hair	·65	Speed (R.V.)	·86	Province	·86		
Salutation	·76	Order	·62	Thought Production	·84	Name	·83		
Departure	·73	Cleanliness of Dress	·60	Neologisms	·15	Duration of Stay	·82		
Handshake	·73	Shaving	·52	Delusions	·01	Date	·78		
Motor Speed	·72	Smell	·34						
Mannerisms	·61	Wetness	·32						
Degraded	·32								

Insight and Perception are not included as they represent single items.

4. Test-retest reliability coefficients were calculated to determine the degree of consistency in the obtained ratings by comparing each rater's initial assessments with the ratings he assigned the same subjects after a sixty-day interval. These results are summarized in Table V.

TABLE V

Test-Retest Reliability Coefficients After 60 Day Period for 20 Controls

	X	Y	Z
Appearance	·51	·67	·72
Behaviour	·80	·92	·89
Mood	·08	·59	·32
Thought	·43	·85	·78
Memory	·92	·89	·89
Hallucinations	·16	·71	·34
Insight	·70	·71	·54

DISCUSSION OF RESULTS REGARDING RELIABILITY AND VALIDITY

Donnelly's results (see Table I) indicate that ratings of individual differences in mood do not accord well with variation found in other behavioural categories. While to a certain extent this lack of relationship would result from instability of mood in the individuals in the sample, the correlations between mood and the other categories were so low that an analysis of the mood items was undertaken to find the source of the unreliability.

An examination of the distribution of item scores indicated that in the case of mood appropriateness and mood tempo, five-point scaling was a spurious refinement since the distributions were basically dichotomous. The distributions of ratings of dominant mood did not differ from the normal. Despite the distributions involved, inter-item correlations were calculated to obtain a rough estimate of the relationship between the items. This analysis showed that the grouping of the items mood appropriateness, dominant mood and mood tempo in the same category was not justified. There was a low negative relationship between dominant mood and mood appropriateness, −·18; and a low negative relationship between dominant mood and mood tempo, −·26. The relationship between mood tempo and mood appropriateness was low, but positive, ·41. The combination of these heterogeneous variables contributed heavily to the extreme unreliability of the mood category total score.

In view of the non-normal distribution of mood appropriateness and mood tempo, these variables were not further utilized in their original form. The latter was used in combination with other variables in a manner described in detail at a later point in the discussion.

Attention was therefore centred upon the variable dominant mood. Inter-rater coefficients calculated for this item proved to be higher (·52; ·62; ·35) than those obtained for the mood category as a whole and presented in Table III (·05; −·12; ·25). Test-retest reliability coefficients, with a sixty-day interval, were ·69; ·19; ·59: whereas for the original mood category scores they had been as shown in Table V, ·32; ·08; ·59. These reliability coefficients leave much to be desired. It would appear, however, that this shortcoming is but a reflection of the difficulties encountered in any attempt at assessment in the general area of mood. Not only instability of mood in the rated individuals but also variations in the level of rapport established by the rating psychiatrists contribute to the difficulty of obtaining reliable evaluations.

Of greater interest is the fact that the presence or absence of hallucinations appears to be related to variations in the ratings of behaviour, thought, memory, appearance, insight or mood. This apparently independent status of hallucinations leads to interesting speculation regarding the production of hallucinations in normals through stimulus deprivation or through administration of psychotomimetic drugs and hints at a rationale for the selective effects of chlorpromazine, which reduces anxiety but does not affect the tendency to hallucinate. However, unreliability of measurement, particularly as regards mood, could produce this apparent lack of relationship with other aspects of the chronic schizophrenic syndrome. This point will be further considered in the light of the results of a factorial analysis of the data.

In summarizing the data on reliability it seems worth while to comment on the fact that the inter-rater coefficients as presented in Table II range between ·70 and ·80. While one would have hoped for greater inter-rater agreement, it should be remembered that these data include ratings of mood and perception, which would effect a general decline in coefficients. In part, too, these borogove-like coefficients are a function of differences in the clinical approach and the general frame of reference of the three participating psychiatrists. These differences, and the degree to which they are affecting inter-rater reliability, become apparent when one refers to Table III. From these data it becomes obvious that mood should not be rated in the terms utilized in this study and the results of the analysis of the items in the mood category suggest that only the variable dominant mood should be retained in its present form.

Table IV indicates that certain of the items are unrelated to the other aspects incorporated in the same behavioural category—a situation which would tend to reduce category reliability. This is particularly true of the items dealing with neologisms and with delusions and ideas of reference. These apparently bear no relationship to the other items with which they have been classified. The adequacy of the present item classifications will be further discussed in the light of the results of the factorial study.

The material on test-retest reliability which is presented in Table V is based on twenty of the forty cases only—those comprising the control group who received placebo rather than the experimental medication. The experimental group were excluded in order to rule out individual differences in response to medication. The coefficients show that the raters vary in their consistency of assessment. Here again the categories of mood and hallucinations are the least consistent.

Each of the 37 items was examined to ensure that between rater variation was not an important factor. The item appropriateness of mood was rejected on this basis.

The distributions of each of the remaining thirty-six items were examined. In the case of ten items the results indicated that ratings on a five-point scale were unrealistic. These items—order, smell, wetness, insight, hallucinations, degraded, neologisms, obedience, motor speed and mood tempo—yielded "j"-shaped distributions. Rather than leave this data out as not being amenable to parametric methods, it was decided to combine the items such that the distribution would tend to normalize.

The method of combination involved dichotomizing each item distribution at the score closest to the median of the distribution. A score of one was given to a rating above the median and a score of zero was given to a rating below the median. Sub-median scores for each of the items were found to be: Order 1; Smell 1; Degraded 1; Wetness 1; Insight 1, 2, 3 and 4; Hallucination 1 and 2; Neologisms 1; Obedience 1; Motor Speed 1 and 2; Mood Tempo 1; which were accordingly given a zero score.

A table of co-occurrence was drawn up to show for each case in the sample the pattern of one and zero scores. It was found that seven variables formed one cluster or group while the remaining variables tended to form a second independent cluster. These sets of items were labelled Group A and Group B. Group A includes the items Order, Smell, Wetness, Degraded, Obedience, Speed (motor), and Insight. Group B includes the items Mood Tempo, Neologisms and Hallucinations. Scores were derived for each individual in the sample by simply counting the number of scores that he obtained (e.g., the number of higher than median ratings) in each of Group A and Group B. These groups were found to be slightly negatively related ($r = -\cdot 20$). For Group A the distribution did not differ from the normal. For Group B, however, which comprises only three items, the distribution was positively skewed and could not be normalized. In consequence, correlations involving Group B can be expected to be relatively low. However, this data was of such interest that it was retained in the study.

Rationale for Factor Analysis

In this analysis we were seeking to obtain further understanding of the structure of chronic hospitalized schizophrenia through studying the factorial structure underlying the ratings. We had, in the course of building the scale, been involved in much discussion as to whether we were dealing with a unitary

disorder in which the symptom sub-groups (catatonic, paranoid, etc.) were differing expressions of the same disease process or whether we were dealing with a group of separate disorders having certain symptoms in common. A preliminary decision was made to place emphasis on category, rather than total, scores. This procedure implies the acceptance of the latter rationale; that is, schizophrenia is a generic term which encompasses a group of disease entities which have certain symptoms in common. This rationale may be tested on the basis of the inference that when the ratings of schizophrenia are factor analysed the obtained factor structure would reflect the underlying pattern of separate diseases. Accordingly, we would expect, if clinical classification be valid, to obtain a paranoid, a simple, a hebephrenic and a catatonic factor. This theoretical position can be formalized into the following hypothesis:

When a factor analysis is carried out on ratings derived from the Weyburn Assessment Scale, the factors extracted should present patterns of factor loadings such that they may readily be identified as descriptive of the basic clinical sub-divisions of schizophrenia.

At this point in the discussion many readers will take issue with the authors. Disagreement will centre around two points. Some will take the point of view that whatever the nature of the results they will be without practical importance, while others will feel that the method of factor analysis is unsuited to the problem and to the investigation of the hypothesis under examination.

Of those who see the results as unimportant, some may hold the view that regardless of the outcome of any factor analysis, the nosology derived through decades of clinical experience and observation is unlikely to be modified. Others will point out that in clinical practice there is a trend to regard diagnostic considerations as secondary to those of treatment, since each case, being unique, demands an individualized therapeutic approach—a situation which renders any procedure of fitting the specific case into a broad diagnostic category of little value.

While these arguments are expressive of an extremely practical outlook, it remains evident that while the problem of the nature of schizophrenia remains unsolved, clinical methodology in this area cannot be based on scientific understanding.

Eysenck (1953) has pointed out that, in general, factor analysis is the method best suited to problems of classification—an observation particularly cogent to our present study in which we are seeking to clarify the structure of one specific area of schizophrenia. However, while there may be general agreement on this point, we are open to challenge with regard to the method of factor analysis used in this investigation. Cattell (1956), we believe, summarizes the case against our procedure when he suggests that the general schizophrenia factor should have been obtained as a second order factor, thus leaving the factor more determined.

Although the method that Cattell proposes has a great deal to recommend it, in the authors' view, there are certain considerations which contra-indicate its use.

Clinical experience with diagnostic classifications will tend to exert an influence upon ratings made by psychiatrically trained people. The presence of one symptom of a symptom complex will enhance the probability that the other symptoms of that particular syndrome will be observed in the same individual; not only because such symptoms apparently tend to co-occur, but also because the psychiatrist expects and looks for them to do so. This is particularly evident in those studies in which the variables involved are

symptoms. By using actual aspects of behaviour, rather than symptoms as such, we have attempted to minimize this difficulty.

Guertin (1951) refers to this halo effect, yet proceeds to enhance its reflection in the data by using Thurstone's multiple group centroid method of factor analysis, "with only the symptoms showing tightest clusterings determining the centroids". When this method is used, the probability exists, even when the ratings are not on a specific symptom basis, that the text-book is being transposed through the data into the obtained factor structure, and that the outcome has been pre-determined by the method employed.

If what we call schizophrenia is, in fact, a group of diseases, the authors believe that the only way to isolate these differences is to remove the communality which will be accounted for by a general factor and to examine the nature of the remaining group factors, for in these should lie the key to the nature of the subsumed disease entities. If such entities exist within the rather narrow confines of the study, that is, within the chronic hospitalized schizophrenia group, the nature of the data should enhance the probability of their being readily demonstrated. If they are distinctive disorders their differences should be indicated by a lack of correlation in some areas and clusters of correlation elsewhere. This would tend to maximize the group factors as opposed to the general factor. In this situation there is a high level of inter-relationship between the variables as we are dealing with the interior of the syndrome. Thus there is for each factor a minimum of material in the hyperplane. This relative density of the matrix will enhance the tendency for intra-syndrome group factors to appear. It therefore follows that if the procedure of breaking chronic schizophrenia down into various clinical sub-types has any validity, these classifications should be readily identifiable in the factor structure and should be clearly demonstrated since they can be expected to account for a relatively large amount of the variance.

Variables Included for Analysis

The memory scale items were highly inter-correlated as shown in Table VI. Because of this high degree of inter-relationship, it was deemed inadvisable

Table VI

Intercorrelations of Memory Items

	Own Name	Hospital	Ward	Date	Province	Duration of Stay
Own Name	—					
Hospital	·70	—				
Ward	·79	·81	—			
Date	·75	·71	·69	—		
Province	·70	·70	·79	·66	—	
Duration of Stay ..	·56	·64	·64	·80	·60	—
Birthday and Age ..	·68	·75	·80	·74	·77	·77

to include them as a group. A preliminary factorization of the memory items was therefore carried out to determine the extent to which they were univariate. The Thurstone centroid method of analysis as demonstrated in Thurstone (1947) was used. The first factor extracted accounted for 73 per cent. of the variance. The loadings obtained were: Name ·83; Hospital ·86; Ward ·89; Date ·86; Province ·84; Duration of Stay ·81; Birthday and Age ·89.

A second factor was extracted but accounted for only ·5 per cent. of the

TABLE VII

	Hair	Shaving	Clean (Hands and Face)	Nails	Shoes	Cleanliness (dress)	Salutation	Handshake	Posture	Mannerisms	Departure	Rapport	Gait	Dominant mood	Speed (R.V.)	Sentence Structure	Delusions	Associative Processes	Thought Production	Thought Blocking	Ward	Group A	Group B
Hair		.19	.67	.34	.33	.23	.62	.45	.23	.03	.51	.55	.45	.43	.72	.20	.06	.66	.79	.67	.64	.20	
Shaving	.39		.36	.53	.39	.09	.51	.05	.50	.24	.46	.61	.56	.37	.24	.73	.27	.72	.62	.57	.24		
Clean(handsandface)	.35	.17		.15	.36	.25	.14	.58	.54	.20	.24	.70	.51	.27	.65	.71	.27	.67	.62	.17			
Nails	.31	.19	.36		.48	.04	.56	.48	.60	.07	.24	.72	.28	.31	.81	.73	.18	.55	.19				
Shoes	.44	.19	.42	.02		.15	.55	.47	.41	.23	.50	.26	.47	.52	.77	.67	.17	.04					
Cleanliness (dress)	.36	.22	.21	.16	.27		.44	.39	.60	.34	.43	.61	.54	.40	.68	.62	.42						
Salutation	.41	.23	.24	.12	.42	.20		.49	.54	.01	.14	.72	.55	.32	.66	.10							
Handshake	.33	.20	.18	.12	.52	.32	.48		.47	.39	.38	.26	.28	.44	.26								
Posture	.44	.35	.10	.24	.32	.19	.39	.33		.31	.45	.61	.54	.24									
Mannerisms	.22	.18	.23	.29	.45	.23	.53	.22	.21		.53	.72	.55	.31									
Departure	.32	.25	.26	.17	.42	.24	.48	.36	.48	.39		.77	.45	.40									
Rapport	.46	.19	.31	.24	.52	.32	.14	.45	.59	.29	.41		.65	.32									
Gait	.50	.25	.19	.29	.32	.19	.44	.47	.54	.28	.48	.67		.44									
Dominant mood	.25	.12	.27	.17	.45	.23	.39	.39	.47	.27	.48	.67	.25										
Speed (R.V.)	.45	.21	.34	.24	.42	.24	.53	.49	.54	.23	.50	.70	.51	.37									
Sentence Structure	.49	.24	.34	.33	.42	.24	.48	.33	.54	.34	.43	.72	.56	.27	.24								
Delusions	.17	.08	.11	.06	.16	.15	.14	.22	.21	.01	.14	.26	.28	.31	.65	.20							
Associative Processes	.44	.21	.32	.27	.33	.18	.44	.36	.48	.39	.38	.61	.47	.52	.81	.73	.27						
Thought Production	.40	.18	.25	.22	.37	.22	.49	.45	.59	.31	.45	.72	.54	.40	.77	.71	.27	.72					
Thought Blocking	.49	.23	.30	.26	.25	.20	.53	.47	.54	.29	.53	.77	.55	.52	.68	.67	.18	.67	.79				
Ward	.38	.16	.34	.27	.39	.23	.50	.46	.47	.28	.41	.60	.45	.32	.66	.62	.17	.55	.62	.57			
Group A	.51	.29	.38	.29	.55	.34	.58	.49	.38	.27	.48	.67	.65	.44	.24	.10	.42	.04	.19	.17	.24		
Group B	.08	.09	.05	.09	.19	.20	.24	.20	.25	.11	.14	.18	.25	.24	.26	.10	.06	.66	.19	.17	.24	.20	

https://doi.org/10.1192/bjp.104.435.359 Published online by Cambridge University Press

variance. The memory items thus seemed clearly univariate and the item Ward was included in further calculations as representing the memory variable since it had the highest first centroid loading.

Twenty-three variables were included in the final matrix. Of these, twenty-one were included in their original form. The other two variables represent the total scores for each individual of the items which were dichotomized and combined to make Groups A and B.

The matrix of intercorrelations was found between the variables or scale items and is presented in Table VII.

Results of the Factor Analysis

Six factors were extracted leaving insignificant residuals. The centroid loadings on these factors are presented in Table VIII.

Table VIII

Centroid Loadings	1	2	3	4	5	6
1. Hair	·62	·21	−·12	·17	−·05	·14
2. Shaving	·35	·17	−·11	·25	·04	·23
3. Clean (Hands and Face)	·48	·45	−·26	−·25	·25	−·05
4. Nails	·43	·51	−·24	−·39	·15	·08
5. Shoes	·60	·18	−·19	·14	−·23	−·13
6. Cleanliness (Dress)	·42	·32	−·35	−·21	·03	·10
7. Salutation	·67	−·19	−·09	·23	·18	−·08
8. Handshake	·60	−·33	−·16	·26	·25	−·20
9. Posture	·70	−·18	−·11	·18	−·13	·26
10. Mannerisms	·31	·25	·28	·18	−·18	−·07
11. Departure	·59	−·25	−·07	·23	·31	·05
12. Rapport	·81	−·18	·21	−·07	−·09	·13
13. Gait	·74	−·05	−·15	·13	−·16	·12
14. Dominant Mood	·55	−·25	−·05	−·13	−·15	·06
15. Speed (R.V.)	·82	−·15	·25	−·05	·06	·02
16. Sentence Structure	·79	·10	·34	−·04	−·06	·06
17. Delusions	−·32	·27	·20	·29	·22	·11
18. Associative Process	·71	·15	·43	·09	·07	·02
19. Thought Production	·80	−·18	·29	−·09	·06	·17
20. Thought Blocking	·81	−·16	·33	−·06	·16	·15
21. Ward	·74	·04	·27	−·03	·11	−·22
22. Group A	·79	·09	·03	·04	·04	·07
23. Group B	−·28	·32	·26	·31	·12	·21
Percentage of Variance	40	6	5	4	2	

Simple structure was obtained after nine rotations. The rotated factor loadings are presented in Table IX.

Discussion of Results

The first factor, which accounted for some 40 per cent. of the variance, is clearly a general factor upon which all variables are relatively highly loaded. In the case of delusions and Group B (viz. mood tempo; neologisms, obscenities and puns, and hallucinations) the loadings are negative. The inference might be drawn from these results that in so far as delusions and hallucinations may be more numerous and more intense in acute schizophrenia than in chronic schizophrenia, this particular variable may work in a reverse direction to the other measures descriptive of the chronic syndrome. Some support for this view may be found in the work of Miller and his associates (1953) who found in evaluating progress in patients suffering from chronic schizophrenia that a group in treatment who show less disturbance in behaviour and improved rapport show more hallucinations than a control group. Possibly an alternative

TABLE IX

Rotated Factor Loadings	1a	2c	3c	4d	5d	6c
1. Hair	·60	·20	−·07	·16	−·12	·22
2. Shaving	·34	·30	·04	·22	−·08	·15
3. Clean (Hands and Face)	·37	·01	·04	−·15	−·13	·67
4. Nails	·30	−·02	−·09	−·08	−·01	·76
5. Shoes	·59	·01	−·19	−·13	−·32	−·08
6. Cleanliness (Dress)	·34	−·05	−·05	·11	−·10	·55
7. Salutation	·69	·03	·30	−·01	−·13	−·04
8. Handshake	·63	−·07	·45	−·03	−·21	−·13
9. Posture	·73	·04	·03	·26	·05	−·03
10. Mannerisms	·32	·28	−·31	−·17	−·08	−·06
11. Departure	·61	·09	·43	·05	·00	−·05
12. Rapport	·83	−·03	−·07	−·03	·26	−·05
13. Gait	·75	·00	−·05	·24	−·08	·07
14. Dominant Mood	·57	−·25	−·02	·11	·13	−·02
15. Speed (R.V.)	·83	·01	·04	−·17	·21	−·03
16. Sentence Structure	·78	·16	−·19	−·20	·19	·07
17. Delusions	−·33	·49	·07	−·09	−·02	−·02
18. Associative Process	·70	·31	−·11	−·30	·16	·01
19. Thought Production	·81	·05	·03	−·09	·36	−·02
20. Thought Blocking	·82	·10	·11	−·17	·37	·09
21. Ward	·73	·05	−·01	−·39	·01	·05
22. Group A	·77	·12	−·01	−·03	·02	·19
23. Group B	−·28	·56	−·06	−·02	·03	−·03
Percentage of Variance	39	4	3	3	3	2

explanation would be that as patients develop better contact with staff members, the staff members become more readily aware of the patient's hallucinatory behaviour or delusional thinking—symptoms which the patient, as he improves, is more likely to discuss. Another possibility may be that though we have a syndrome which shows a high level of internal consistency, as demonstrated by the degree of inter-correlation between scale items and by the relatively heavy saturation of every item with the general factor, we may not be measuring chronic schizophrenia as such but the deterioration effects of chronic mental illness and long-term hospitalization which may not be specific to schizophrenia.

If this were the case, the fact that hallucinations and delusional thinking correlate negatively with the rest of the syndrome could suggest that the chronic pattern as operationally defined in this scale may not be a function of schizophrenia. Rather, it might be related to the level of hospital care—to a deteriorative process resulting from living for prolonged periods of time in a "back ward" environment.

Unfortunately, the question of specificity of the results of this investigation to schizophrenia is outside the scope of the study. It cannot be answered from the collected data. However, with regard to the problem under examination, it can be stated to the extent that the ratees are schizophrenic, the nature of this disorder should be reflected, as hypothesized, in the obtained factor structure.

As has been stated, the general factor accounts for some 40 per cent. of the variance; the remaining factors together account for something less than 20 per cent. of the individual variation.

The highest loading variables on factor two are Group B (Mood Tempo, Neologisms, and Hallucinations) ·56; Delusions ·49; and Associative Processes ·31. This factor appears to be associated with thought disorder.

On factor three, the variables showing the highest loadings are Handshake ·45; Departure ·43; Salutation ·30; and loading negatively Mannerisms −·31. This factor appears to be associated with the patient's ability to establish and

to discontinue contact with the psychiatrist; that is to say, his approach to other people.

Factor four is ill-defined. However, the pattern of loadings suggests that it is a motor factor.

Factor five loads most highly on Thought Blocking ·37 and Thought Production ·36. This factor is also difficult to identify, but it appears to be associated with an unwillingness to speak.

Factor six shows high loadings on Cleanliness of Nails ·76; Cleanliness of Hands and face ·67; and Cleanliness of Dress ·55 and is clearly a factor of general cleanliness, which may be related to nursing care.

The results of the factorization do not support the hypothesis that the analysis of ratings would give rise to factors presenting patterns of factor loadings such that they might readily be identified as descriptive of the basic clinical sub-divisions of schizophrenia.

The results of this orthogonal analysis yield strong evidence that chronic hospitalized schizophrenia is essentially a unitary phenomenon.

W.A.S. Second Revision

The data presented in this report form the basis of the second revision of the Weyburn Assessment Scale.

The revised scale is intended specifically as an instrument for the assessment of the degrees of illness in chronic hospitalized schizophrenic patients. For this group, it will provide a relatively reliable measure of change of behaviour and will offer an objective assessment of the success or failure of therapeutic procedures.

Bibliography

Baker, A. A., and Thorpe, J. G., "Deteriorated Psychotic Patients—Their Treatment and Its Assessment", *J. Ment. Sci.*, 1956, **102,** 780–789.

Cattell, R. B., Personal communication, 1956.

Cohen, L. H., Malmo, R. B., and Thale, T., "Measurement of Chronic Psychotic Overactivity by the Norwich Rating Scale", *J. Gen. Psychol.*, 1944, **30,** 65–74.

Donnelly, J., "Clinical Trial of Chronic Schizophrenia Patients Under Nucleotide Therapy", 1953. Unpublished Report.

Eysenck, H. J., *The Structure of Human Personality*, 1953. London: Methuen.

Guertin, W. H., "A Factor Analytic Study of Schizophrenic Symptoms", 1951. Thesis, Univ. of Mich., University Microfilms, Ann Arbor.

Idem, "A Factor Analysis of Schizophrenics Rated on the Activity Rating Scale", *J. Clin. Psychol.*, 1956, **12,** 163–166.

Lorr, M., "Multi-dimensional Scale for Rating Psychiatric Patients", 1953. Veterans Administration Technical Bulletin, T.B. 10–507, Veterans Administration, Washington.

Lucy, J., Clancey, J., Hoffer, A., Osmond, H., Smythies, J., and Stefaniuk, W. B., "Design and Planning in Psychiatric Research as Illustrated by the Weyburn Chronic Nucleotide Project", *Bull. Menn. Clinic.*, 1954, **18,** 147.

Malamud, W., and Sands, S. L., "A Revision of the Psychiatric Rating Scale", *Amer. J. Psychiat.*, 1947, **104,** 231–237.

Miller, D., Clancy, J., and Cumming, E., "A Method of Evaluating Progress in Patients Suffering from Chronic Schizophrenia", *Psychiat. Quart.*, 1953, **27,** 439–451.

Plant, J. S., "Rating Scheme for Conduct", *Amer. J. Psychiat.*, 1922, **1,** 547–572.

Scherer, I. W., "A Behaviour Rating Scale for use in Activity Therapy Situations", Info. Bull., Dept. Med. and Surg., Psychiatric and Neurological Div., The Veterans Administration, January, 1951.

Thurstone, L. L., *Multiple Factor Analysis*, 1947. University of Chicago Press.

Wilcox, P. H., "The Gardener Behaviour Chart", *Amer. J. Psychiat.*, 1942, **98,** 874–880.

Whittenborn, J. R., and Holzberg, J. D., "The Generality of Psychiatric Syndromes", *J. Consult. Psychol.*, 1951, **15,** 372–380.

Wittman, P. M., "The Elgin Check List of Fundamental Psychotic Behaviour Reactions" *Amer. Psychol.*, 1948, **3,** 280.