The Development of the Blackfoot Clinical Rating Scale for Evaluating and Recording Personality Changes in Mentally Ill Patients

John R. Cochran
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Psychology Commons

Recommended Citation
https://digitalcommons.usu.edu/etd/6399

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact dylan.burns@usu.edu.
THE DEVELOPMENT OF THE BLACKFOOT CLINICAL RATING SCALE FOR EVALUATING AND RECORDING PERSONALITY
CHANGES IN MENTALLY ILL PATIENTS

by

John R. Cochran

A thesis submitted in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE
in
PSYCHOLOGY
1950

UTAH STATE AGRICULTURAL COLLEGE
Logan, Utah
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Purpose</td>
<td>1</td>
</tr>
<tr>
<td>Scope</td>
<td>2</td>
</tr>
<tr>
<td>Importance</td>
<td>2</td>
</tr>
<tr>
<td>Review of Literature</td>
<td>4</td>
</tr>
<tr>
<td>Rating Scales</td>
<td>4</td>
</tr>
<tr>
<td>Definitions and Purposes</td>
<td>4</td>
</tr>
<tr>
<td>Types</td>
<td>4</td>
</tr>
<tr>
<td>Reliability</td>
<td>7</td>
</tr>
<tr>
<td>Definitions</td>
<td>7</td>
</tr>
<tr>
<td>Methods of Establishing</td>
<td>8</td>
</tr>
<tr>
<td>Acceptability of Methods Used</td>
<td>10</td>
</tr>
<tr>
<td>Validity</td>
<td>11</td>
</tr>
<tr>
<td>Definitions</td>
<td>11</td>
</tr>
<tr>
<td>Methods of Establishing</td>
<td>12</td>
</tr>
<tr>
<td>Acceptability of Methods Used</td>
<td>14</td>
</tr>
<tr>
<td>Method of Procedure</td>
<td>18</td>
</tr>
<tr>
<td>Development of the Blackfoot Clinical Rating Scale</td>
<td>18</td>
</tr>
<tr>
<td>Establishment of Reliability and Validity</td>
<td>21</td>
</tr>
<tr>
<td>Blackfoot Clinical Rating Scale</td>
<td>23</td>
</tr>
<tr>
<td>Definitions</td>
<td>23</td>
</tr>
<tr>
<td>Results</td>
<td>38</td>
</tr>
<tr>
<td>Conclusions</td>
<td>42</td>
</tr>
<tr>
<td>Summary</td>
<td>43</td>
</tr>
<tr>
<td>Bibliography</td>
<td>45</td>
</tr>
</tbody>
</table>
INTRODUCTION

Purpose

The problem discussed in this thesis is the conception, development, and application of the Blackfoot Clinical Rating Scale, together with methods employed in establishing reliability and validity.

Understanding the personality structure of an individual is one of the most difficult tasks anyone can undertake. The desire to make this understanding easier and to structure an approach which would be helpful led to the development of the Blackfoot Clinical Rating Scale.

It was the opinion of the staff at State Hospital South at Blackfoot, Idaho that there was no device available adequate for recording clinical impressions. Since the object was to provide a means of noting clinical impressions gained over a period of many days of contact with a patient, and doing so as quickly as possible, the rating scale method was chosen. The Blackfoot Clinical Rating Scale was constructed to answer the need for a convenient device for recording clinical impressions of the personality development of mental patients in a consistent and objective manner. This scale will be referred to, henceforth, as the Blackfoot Clinical Rating Scale, since it was developed at State Hospital South in Blackfoot, Idaho. It is designed as a basic psychometric tool to be used chiefly by clinical psychologists,
psychiatrists, and others working in mental hospitals.

Scope

It would be impossible and unnecessary to review here all of the behavior disorders, personality testing devices, and statistical methods. This thesis will, however, serve several important functions. A brief discussion has already been given of how the Blackfoot Clinical Rating Scale was conceived and prepared to meet a definite need. Another section presents the manner in which the individual items or qualities on the scale were chosen, and their definitions prepared. In this same section, support is given for the choice of a six-point scale rather than one of a different number of scoring points. The use of the definitions with the scale is explained. Two major sections present the methods used in establishing validity and reliability. Derivation of validity and reliability of the Blackfoot Clinical Rating Scale is explained. The scale has been evaluated by the author and by staff members of the hospital. Conclusions are then drawn and a summary is made of the work done.

Importance

Most hospitals have some device by which they attempt to appraise a patient's mental and emotional condition at various times. However, in the opinion of the clinical staff of State Hospital South, no satisfactory device has thus far been prepared for which validity and reliability have been adequately established. It would be of
considerable benefit to any clinician to have a personality rating scale which would give honest and consistent results. The Blackfoot Clinical Rating Scale is believed to be the first of its kind which is valid and reliable.

Rogers (26) has observed that diagnostic thinking requires a framework if one is to cover the many aspects of the personality thoroughly and omit no important ones. Staff members of State Hospital South agree that the Blackfoot Clinical Rating Scale serves very well as this necessary framework.

By virtue of its carefully selected and defined terms, this 134-item, 6-point scale has the advantage of indicating either regression, improvement, or no gain under therapy. This sensitivity to personality changes provides an objective evaluation of the patient at various times so that changes in therapy may be made if necessary.
REVIEW OF LITERATURE

Rating Scales

Definitions and purposes. In order to adequately define rating scales, it is necessary to describe their purposes. For this reason, the definitions and purposes of rating scales are considered together in this section.

According to Shaffer (26), a rating scale "is a device for systematizing and improving the expression of opinion concerning a personality trait." Such (27) defines a rating scale as "a device for defining the trait under question and grading the individual on the basis of this trait."

Rogers (26) has observed that diagnostic thinking needs a framework if one is to cover the many aspects of the personality thoroughly and omit no significant ones. He constructed a rating scale for problem children using the component factor method of analysis. The purpose of the scale was not to obtain a score value, but to be an aid to an objective type of analytic thinking. In his opinion, when the ratings on the various traits of an individual have been completed, the scale serves as an aid in planning treatment. The Black-foot Clinical Rating Scale fills this same dual purpose by providing a framework for analytic thinking and for planning therapy.

Hunt (13) says a rating scale is "a means of indicating
quantitatively the degree to which individuals possess abilities or traits. In this method, however, instead of having the individual perform some task or problem which is to indicate the amount of the trait, the measurement represents the subjective impressions of someone who judges the amount of the trait or ability possessed by the individual from previous associations with him in situations where the trait or ability is supposed to be shown. Patterson (23) has the following to say in defense of rating scales and subjective impressions. "Most criteria are founded on human judgment. Thus, we should get the best possible judgments by providing competent trained observers with the means to render sound systematic judgments." These judgments "can be most accurately and conveniently expressed by use of a rating scale technique." A rating scale "enables one to see the specific uncontrollable elements and appraise the total performance."

**Types.** Five fundamental types of rating scales are described by Such (27). In the first, the method of paired comparisons, "the judge compares every individual with every other in the group of subjects to be rated, and in each pair rates one as superior to the other in the trait under consideration." The subject who receives the greatest number of "firsts" in the comparisons gets the highest score. The score is a relative position, not an absolute score. Naturally, the number of judgments increases much more quickly than the number of cases, making this method impracticable for large groups. Since the Blackfoot Clinical Rating Scale
was designed for use with unlimited numbers of persons, this first method could not be employed.

The second, the order-of-merit method, requires that the subjects be lined up in a 1, 2, 3, etc., order by choosing the best and then the next best until all have been ranked. Considerable difficulty is encountered in attempting to keep in mind the whole field and each individual until the best is chosen. In this method again, the subject receives a relative position and not an absolute rating.

The third, the absolute rating scale method, permits the judge to give an absolute value to each trait. Only one judgment is necessary for each trait, so this method is more rapid than the two preceding ones. The Blackfoot Clinical Rating Scale constitutes an elaborated application of this method in that both an absolute descriptive statement and a numerical value are given for each rating under each trait.

In the fourth, the man-to-man method, the rater equates the rated person to some known standard persons who have been selected by pooling the opinions of several judges. The problem of setting up standard persons and of not having the objective definitions made this method unsuitable as a pattern for the Blackfoot Clinical Rating Scale.

The fifth, the check-list method, simply requires the judge to read a list of traits and check those which apply to a given person. The Blackfoot Clinical Rating Scale was not based on this method, because a check-list would have furnished very little information of the kind desired. The
well-defined gradations of personality traits would not have accompanied a simple check-list.

Additional lists of the kinds of rating scales have been described by others, but they do not differ enough from Such's list to make a discussion of them of value here.

Investigators have used a variety of methods of assigning values to their ratings. Some have preferred to indicate quantities of traits by using minus and plus values. Some have simply used a continuum on which the rating can be checked anywhere between the two extremes and may or may not carry a numerical score. Kelly (11) used fifteen points under each trait on his rating scale. However, he found it too difficult and confusing to construct sufficiently finely graded definitions for so many points under each term, so he merely described three points, the two extremes and the center.

Ratings of "0" to "5" were used for the Blackfoot Clinical Rating Scale because this number of divisions and breadth of distribution permits a patient's degree of any trait to be analyzed sufficiently to be meaningful and yet not be cumbersome.

Reliability Definitions. Writers on the subject appear to be in close agreement in their definitions of reliability. Greene (10) defines reliability in terms of the degree to which a test gives consistent results. Smith (30) describes reliability as the self-consistency of a test. He would have the results
of a test be stable and predictable. If a test has a good
degree of reliability, those who get high scores with a test
one time should get high scores a second time. The rank-
order of scores obtained from a second administration of a
test should be quite similar to that obtained from the first
administration if the test is to differentiate between indi-
viduals. A rating scale that has high reliability should
permit several judges to rate independently any given subject
at a given time and to obtain a high correlation between
their ratings.

Methods of establishing. The accepted methods of estab-
lishing reliability usually involve the use of one of three
reliability coefficients designated by the symbol "r." Ac-

According to Greene (10) and Smith (30), these coefficients
are of retest reliability, split-half reliability, and

equivalent-form reliability. Retest reliability is obtained
by correlating the scores from one administration of the test
to a group with the scores from a later administration of the
same test to the same group. Split-half reliability is
usually accomplished by correlating the odd-numbered items
of a test with the even-numbered items. Equivalent-form

reliability is found by correlating one form of a test with
another equivalent form.

According to Rush (27), the reliability of a rating
scale is usually determined by finding out how well two sets
of judges, using the scale, agree on the traits of the same
individuals. This method applies the retest process, except
that, here the subject, instead of taking a test twice, is given two or more ratings by different judges.

Symonds (31) has reported that reliability has been established for some rating scales by having two raters rate the same children. He mentions that this method was used for the Personality Rating Scale of Harston on which children are rated for introversion and extroversion.

Rogers (26) made a preliminary study of the reliability of his seven-point rating scale by having six raters rate each of five cases on each of the seven factors. Of the total of 210 ratings, on 66 percent only was there agreement within two categories on the seven-point scale.

Kelly (14) established reliability on his 36-trait personality rating scale with comparisons of the independent ratings of judges.

Cohen, Malmo, and Thale (3) have written on the reliability of the Norwich Rating Scale, which is used to rate such factors as over-activity and untidiness in psychotic patients. They obtained correlation coefficients of .65 to .90 for separate traits and .81 to .82 for composite scores. Two simultaneous, independent ratings of ten patients yielded these coefficients.

Schney (29) has reported on the reliability of the Wilke Personality Rating Scale. The coefficients of correlation between paired raters ranged between .57 and .70. Computing from these correlations, they found reliability coefficients of .73 to .82 for the total ratings.
Acceptability of methods used. The principal method used in establishing the reliability of the Blackfoot Clinical Rating Scale was that of determining the correlations between the independent ratings of three judges on the traits of the same individuals. Substantial support for this method was reported in the preceding section from Bush (27), Symonds (31), Rogers (26), and Kelly (14). Tables 1 and 2 reveal the very significant correlations which were obtained both on the traits and on the total scores. The preceding section discussed some of the investigations of reliability which have been made on various rating scales. Tables 1 and 2 show that the correlations obtained for the Blackfoot Clinical Rating Scale compare quite favorably with those for the Norwich Rating Scale, and are considerably more significant than those for the Wilke Personality Rating Scale. They are also much more significant than the percentages of agreement which Rogers (26) achieved. Rogers explained his lowest agreement, that on self-insight, by suggesting that this term is probably the most intangible and the least adequately defined. The higher correlations on the Blackfoot Clinical Rating Scale are probably due in large part to the consistent efforts to define each term and each rating so clearly that, even though intangible, they would be thoroughly understood by any qualified judge.

Since the Blackfoot Clinical Rating Scale is designed to be sensitive to small changes in behavior, its reliability could not be based on any method of repeated ratings of the
same subjects after intervals of time. This sensitivity requires that changes in ratings must occur in response to improvement or deterioration of patients.

Validity

Definitions. Some difference of opinion exists regarding even the general definition of validity. Validity is a concept often difficult both to bound and to demonstrate.

In common usage, validity is the degree to which a test actually measures what it purports to measure. English (7) has defined validity as the "extent to which a test measures the trait or ability intended." Cronbach (4) insists that a modification of the usual concept of validity is necessary. He says that a test is valid "to the degree that we know what it measures or predicts." At first glance, the change of emphasis may seem trivial, but a more deliberate consideration reveals that this idea is significant. The direction or the nature of the approach to the analysis of the validity of a test is of consequence.

Bowers (1), in an even more general view of validity in science, states that if a proposition is consistent with other related propositions it is valid in that system of propositions. "Validity is logical consistency within a conceptual matrix, and has nothing to do with empirical truth. The empirical truth probability of the proposition is determined by research verification."

In the following pages, the methods are described which, it is believed, will show that the Blackfoot Clinical Rating
Scale is logical and consistent. Research is discussed which verifies the empirical truth of the proposition that the Blackfoot Clinical Rating Scale gives an adequate picture of the personality of a mentally ill patient.

Methods of establishing. There are several avenues of approach to the problem of establishing validity. Similarities between the approaches of the different investigators are easily discovered. However, to insure a full and accurate understanding of this subject, each approach must be considered briefly.

In general, validity is usually established by use of the validity coefficient, that is, the coefficient of correlation between the test and some outside measure or criterion of the trait. As Smith (30) points out, a major difficulty encountered at this stage is finding a suitable outside criterion. The criterion may vary from one test to another, but the method remains basically the same.

Cronbach (4) describes two basic approaches to validity: logical analysis and empirical analysis. In logical analysis, one tries to determine precisely what the test measures. In empirical analysis, one endeavors to show that the test is correlated with some other variable, and therefore measures the same element. The former leads to a psychological characterization of the test. The latter generally relates the test to a practical purpose. Cronbach (4) goes on to explain that validity may be established either deductively by demonstrating that the test conforms to the
definition of the trait to be measured, or inductively by
listing the traits illustrated in the items at hand.

For Greene (10), the term "validity" has two some-
what different meanings, one qualitative and the other quan-
titative. The former requires that a valid test must meas-
ure only what it is intended to measure. The test must
sample all that is relevant and exclude all that is irrele-
vant. The final decision must be based upon some procedure
which limits relevance or validity to the accepted operations.
The qualitative validity of an appraisal is indicated by its
objectivity. As applied to rating scales, the term "objec-
tivity" means that the appraisal is not distorted by the
idiosyncrasies of the appraiser. The opposite is true for
subjectivity. A test is considered free from subjective
distortion if the judges involved agree in their judgments
and in the credits given for each trait. Such (27) has
stated that validity depends on how well the judges under-
stand the definition of the trait and agree in their judg-
ments.

Quantitative validity is established when a test can
be shown to be a very accurate measure of what it professes
to measure. The accuracy of the measure usually is shown
by correlating the test with some outside criterion. When
no outside criteria are available, Smith (30) suggests that
the criterion of internal consistency be used. A large
number of test items is given in a preliminary test; then,
only those items which clearly distinguish between high
scores and low scores are retained.

Weinland (33) has a somewhat different process for finding external criteria. He has studied the construction of rating scales and personal inventories in parallel to measure the same persons, in the same elements, by different methods. The two instruments thus serve to some extent as criteria for each other.

Acceptability of methods used. This section discusses the methods used in validating the Blackfoot Clinical Rating Scale. Support for these methods was given in the preceding section. The first method discussed was the general use of the coefficient of validity. Some patients at State Hospital South had been given projective tests. It was hoped that the rating scale results on these patients could be correlated with their test results. Some definite parallels exist between the ratings and the personality or behavior results on the tests. However, computing a statistical correlation would have required first setting up and defining arbitrary ratings for behavior described by the projective tests. This task was set aside for some later date when additional research on this problem can be done more adequately.

Larson (17) partially validated the Blackfoot Clinical Rating Scale, in regard to total scores, as part of a recent study. He found a correlation significant at the five percent level between the total scores, or level of adjustment, of mental patients rated on the Blackfoot Clinical Rating Scale and individual behavior in a group problem-solving
situation. The group situation was arranged so that each subject's participation and behavior could be secretly observed and evaluated. A modified form of Bales' system for rating group interaction was used to rate the amount of interaction in terms of participation and direction for each subject. Patients' total scores on the Blackfoot Clinical Rating Scale were in nearly the same rank order as those on the modified Bales' form. This result suggests that many of the same elements involved in the total adjustment of a patient are recorded in both devices. Thus, the two validate each other.

Cronbach (4) would have validation established by logical analysis or by empirical analysis. Both methods were attempted with the Blackfoot Clinical Rating Scale. Logically, a psychological analysis of the behavior of a patient in various situations was recorded and scored on the Blackfoot Clinical Rating Scale. An attempt was then made to correlate this score with the descriptions of the patient's case history and behavior disorder as recorded in the files by the psychiatrists and social workers. Pintner and Forlaus (25) report use of a similar method of validation where the teacher's descriptions of the personalities of several hundred school children were used as the "consensus of experts" and correlated with a personality inventory.

While the use of this approach reveals some definite agreement between ratings on the Blackfoot Clinical Rating
Scale and the material in the files, the results were not conclusive. The Blackfoot Clinical Rating Scale is designed to be sensitive to small changes in behavior, and most of the patients had changed between the time of the write-ups and the time of the ratings.

The next question is that of qualitative and quantitative validity. The Blackfoot Clinical Rating Scale satisfies the first by reason of the method of its construction. The clear, accurate definitions for all of the traits and for all of the six-point choices under each trait make it relatively easy to include that which is relevant and exclude that which is irrelevant. The objectivity, which indicates qualitative validity, is attained in two ways. First, use of definitions clearly limited into six-point choices forces judges to select definite positions and scores rather than vague points on a continuum. This forced-choice technique avoids distortion by possible idiosyncrasies or response sets of individual judges. Cronbach (4) reports that the forced-choice technique is the only one in common use which seems to be free from response sets. Second, the Blackfoot Clinical Rating Scale may be accepted as free from subjective distortion, because there is close agreement between the independent judgments of the appraisers. Tables 1 and 2 show the high agreement that was obtained. In Such's (27) opinion, the validity of a rating scale depends on how well the judges understand the definitions of traits to be appraised. The judges were
all clinically trained and experienced and understood the
concepts of each item and the item's clinical relationship
to behavior. With this understanding, the definitions as
given were sufficient, and adequate placement and rating
of the patient could be made.

In Louttit's (18) opinion, validity is weakened if
the judges are not adequately acquainted with the subjects
they rate. Fortunately, the twenty-six patients on whom
reliability was established were all well known to the
judges. The high agreement between the independent ratings
would seem to support this observation.
METHOD OF PROCEDURE

Development of the Blackfoot Clinical Rating Scale

The Blackfoot Clinical Rating Scale in its final form is the result of many hours of careful discussion, weighing, defining, accepting and rejecting terms. A list of 14 qualities was finally accepted. Only those which were essential to a full picture of a patient's personality at any given time and which were adapted to use in a rating scale were included. Dr. Martin, Dr. Lake, and several psychology interns cooperated in the discussion and approval of each term until the final list was accepted. (See Fig. 1) Several authors provided a frame of reference for the discussions (2, 5, 7, 8, 9, 11, 19, 21, 26, 34). Terms were sought, evaluated, and selected which would indicate a patient's level of adjustment in many different areas and which, besides serving as a record of personality development, would be a concrete aid in mentally reviewing, visualizing, and describing progress in personality development. The decision was emphasized that the scale be so devised that clinical impressions could be recorded in a relatively objective manner so that regression, improvement, or no gain under therapy would be revealed.

Weightings of "0" to "5" for each item were selected since they best show equal amounts of progress and permit inter-item comparisons. For each item, a rating of "0"
BLACKFOOT CLINICAL RATING SCALE

<table>
<thead>
<tr>
<th>Patient</th>
<th>Date</th>
<th>Rater</th>
</tr>
</thead>
</table>

1. Self-esteem  
2. Motivation  
3. Self-confidence  
4. Insight  
5. Ability to abstract  
6. Self-expression  
7. Social interaction  
8. Affective tolerance  
9. Absence of tension  
10. Empathy  
11. Mood-outlook  
12. Independence  
13. Memory  
14. Spontaneity

Figure 1. Layout of Blackfoot Clinical Rating Scale
indicates an almost complete lack of the trait, while "5" indicates a very superior and unusual amount of the same trait; "2" is slightly below average while "3" is slightly above average. Some argument was raised that a rating of "5" should be a surplus extreme of a trait in the same manner that "0" is a deficiency extreme. In response, it was contended that as long as the general definition of a trait is accepted, the trait itself is a good and desirable quality and must remain so whether possessed in a quantity of 3 or in a very superior quantity of 4 or even 5. For example, the possession of self-esteem, number 1 on the scale, is a necessary element in mental health and no matter how great one's self-esteem becomes, it remains a healthy, important quality. If it becomes a detriment, it is no longer self-esteem; it has become vanity or selfish pride or something else, and must be so designated.

In order to insure uniform understanding of the personality items, each item and its ratings had to be carefully defined. In the process of defining broadly and inclusively these central terms and of defining more narrowly the rating levels, a variety of sources was employed. Ideas and suggestions were drawn from many authors (2, 7, 11, 12, 19, 20, 34) and incorporated into the present set of definitions, a reproduction of which begins on page 22.

Figure 1 illustrates the scale itself upon which ratings are made. For accurate use of the scale, continuous
reference to the definitions is necessary.

Establishment of Reliability and Validity

Reliability was established by having three judges make independent ratings on twenty-six patients at State Hospital South in Blackfoot, Idaho. The judges were a clinical psychologist, a psychiatrist, and an intern psychologist. Three correlation coefficients for each item were computed by correlating the rating scores of the judges in three different ways. Adding the three correlations and dividing by three gave an average correlation for each item. Table 1 shows the high correlations which resulted for items and for averages of items. All of these correlations are well beyond the one percent level of significance. Of added importance is the fact that the ratings of the three judges were within two categories on all of the items on all of the patients.

Highly significant correlations were also computed for total scores. Table 2 shows these correlations. They are not averages of other correlations, but are the correlations between the total scores of each judge compared with the total scores of the other two.

An attempt was made to validate the Blackfoot Clinical Rating Scale by correlating ratings with results on projective tests and with clinical information from the hospital files. Both methods revealed rather definite relationships. However, neither of the methods were adapted to the computation of statistical correlations without considerable additional work on the information from the files and the results
from the projective tests.

Larsen (17) found correlations significant at the five percent level when he compared total scores on the Blackfoot Clinical Rating Scale with results on a modified Bales form. This result partially validated the total scores.

The high correlations between the independent ratings of the three judges were made possible by the clear, accurate graded definitions. These definitions are highly valid descriptions of the personality traits and behavior which the judges rated. The consensus of expert clinicians was obtained in the approval of the definitions. The expert clinicians concurred in their judgment that the definitions clearly and validly defined the personality areas to which they were to be applied.
BLACKFOOT CLINICAL RATING SCALE

Definitions

In order to insure uniform understanding of the personality elements in the Blackfoot Clinical Rating Scale, each element and its ratings are here defined. Each central heading is one of the personality elements on the scale and is numbered to correspond. Immediately under each element appears the broad definition for that element. Then the ratings from "0" to "5" are individually defined. A patient's rating on each personality element is determined by selecting that numerical rating which most adequately describes the amount of that element which he possesses.

1. Self-esteem

   The term self-esteem includes the following concepts: Self-respect, Self-evaluation. The satisfaction with which the patient regards himself. What he really thinks of himself, not the impression he may try to give socially in order to conceal low self-esteem. How he believes he would be assessed by others if they understood him thoroughly.

   0. Feels absolutely worthless most of the time.
   1. Feels inferior to most people. Can establish very few satisfactory aspirations.
2. Not convinced his worth is quite average.
3. Feels he is using his potentialities better than average.
4. Has a very high regard for himself, for his personal worth.
5. Rarely recognizes a flaw or weakness in his standards. Feels almost no insecurity in his values and point of view.

2. Motivation

The term motivation includes the following concepts. That which initiates, sustains, and directs activity, and helps define the worth of the goal. The variety of disequilibriums which he wants to do something about. Feeling the need for action, for improvement, etc. Feeling a need to do something in order to achieve or acquire satisfaction.

0. Stuporous, apathetic.
1. Occasionally carries through or attempts an activity or expresses needs.
2. Pretty fair amount of drive to do well. Interested and active, but somewhat below normal.
3. Slightly more than average versatility and intensity of interests.
4. Usually very active in persistent and varied undertakings. Definite drive to excel.
5. Varied, intense, and sustained goal-directed activity.

3. **Self-confidence**

The term self-confidence includes the following concepts. How well the patient thinks he can do what he sets out to do. How many things he believes he can do. Willingness to attempt new activities. Assurance and reliance on self.

0. Afraid to attempt anything. Distrustful of own ability to a completely handicapping degree.

1. Usually distrusts own ability. Very unsure. Attempts a few simple activities with very limited self-reliance.

2. Slightly less than average willingness to undertake and believe that he can accomplish a variety of tasks.

3. Slightly more than average willingness to undertake and believe that he can accomplish a variety of tasks.


5. Extremely sure of own ability at most tasks. Extremely well-poised and at ease. Willing to enter nearly any situation or attempt nearly any task.
Insight

The term insight includes the following concepts. The degree to which the patient is able to look at himself objectively. Ability to understand relationship between his feelings and attitudes, and to have his feelings and attitudes available to his self-reactions. Degree of self-understanding. Ability to look at and understand his emotional development of feelings and attitudes and their contribution to his own condition.

0. No understanding of self or ability to objectify.
1. Sees very few of the implications and relationships except from his personal bias with his personal emotional slants.
2. Sees some of the implications of his adjutative techniques. Can view himself in a detached manner in some areas, and recognizes the personal nature of his own feelings about these areas.
3. Can see himself in more areas and with more comprehension, and with less distortion by personal emotions.
4. Can objectify very well most of his feelings and attitudes and understand their development. Can look honestly at his needs, motivations, and reasons most of the time.
5. Almost complete self-understanding with nearly perfect ability to examine his personal feelings and attitudes objectively.

5. **Ability to abstract**

The term ability to abstract includes the following concepts. Ability to extract essential characteristics out of a situation. Conceptualizing. (Brain damages cannot.) Ability to see commonality between objects. (How is a banana like an orange? Both fruits.) It is not an acquired mental set, nor a specific aptitude. It is a capacity level of the total personality. (Absence—the patient cannot keep in mind several aspects of a situation at one time, cannot readily grasp the essentials of a given whole, cannot plan his actions ahead in ideational fashion.) Includes in its scope more than the immediately given situation. (Tests which involve sorting reveal lack of abstraction.)


1. Delays responses and considers complexity only briefly and inadequately; can see some elements, but cannot synthesize them or delay his response. Makes a picture or decision
without nearly considering all factors.

2. Sees some relationships and can delay partially, but not enough for a really adequate abstraction. Can see different variables, but not enough of the relationship for a full gestalt.

3. Can consider various aspects of a problem with delay, and integrate adequately a good part of the time. Achieves many good gestalts.

4. Can integrate noticeably well a great many variables which he grasps readily, and to which he can make careful, delayed responses for a really good adjustment.

5. Can calmly consider all relevant variables and can withhold judgment or integration until the situation is appropriately defined.

6. **Self-expression**
   
The term self-expression includes the following concepts. How well the patient can perceive what he wants, let himself do what he wants to do, and get at, reveal, and express feelings and needs. Willingness to accept responsibility for his own attitudes. Ability to act a feeling readily. (Ability to feel kindly or hostile and to perform acts of kindness or hostility readily and overtly.) Freedom from neurotic inhibitions.
0. Completely rigid, constrictive, inarticulate, inactive, and unable to release any real feeling or to act with any freedom. Completely inhibited.

1. Unable to express more than a few simple feelings and recognize a few attitudes.

2. Aware of many attitudes and able to articulate them, but not in a really well-integrated program that is adequate.

3. Somewhat better than average in ability to get at, reveal, and express his feelings.

4. Noticeably free from neurotic inhibition in most cases. Knows what he wants and is able to express his feelings adequately.

5. Is able to recognize and express nearly all his feelings and needs with no neurotic restraint, and in such a manner that they are understandable to others.

7. **Social interaction**

The term social interaction includes the following concepts. The variety and complexity of the social relations he can have with other people, both with individuals and with groups.

0. Unable to interact socially with anyone. Completely insulated. Rejects, or is indifferent to, all social interaction.
1. Able to have a very limited number of very simple social relations with a few people. Is insulated and, or isolated from most contacts, and is very shy, timid, backward, reserved, and poorly poised.

2. Interacts pretty well with people in a somewhat limited group.

3. Little better than average ability to deal well with a variety of people in fairly complex situations.

4. Superior ability to interact easily with most persons in complex social relationships.

5. Able to enter into a great variety of very complex social relations with all kinds of people and groups.

6. Affective tolerance
   The term affective tolerance includes the following concepts: How much emotion the patient can stand. (Some people when they become very angry go into a rage, or even black out. Others, if they feel threatened or embarrassed, become weak and shaky. For others, strong emotion may serve as a healthy drive for action.) How much positive or negative feeling he can experience without confusion. (Exclude psychopaths and some hysterics who simply show lack of responsiveness.)

6. Goes to pieces very easily; goes to emotional extremes readily. Can stand neither praise
nor criticism. Personality disorganizes under certain specific emotional situations. Some patients may deny or repress all emotion because of inability to deal with it.

1. Over-reacts emotionally most of the time. Can make only minimal delayed or integrated reactions to many specific situations. May, in special cases, deny or repress most of his emotion.

2. Can delay and integrate under mild and some strong emotion, but not generally enough to make an adequate adjustment.

3. Can delay and integrate under emotional conditions for a reasonably good adjustment in a particular situation.

4. Under emotional impacts, can deal with his feelings well most of the time. Can handle frustration without loss of control.

5. Almost perfect emotional tolerance or control. Extremely seldom upset or excited by even severe emotional situations. Is aware of and can deal constructively, and with delay, with intense emotion.

9. Absence of tension

The term absence of tension includes the following concepts. Degree of strain, anxiety, or apprehension felt. How well he can relax. Degree of freedom from
unhealthy muscular constriction. Tension is a condition of disequilibrium or dissatisfaction reflected in muscular and visceral reactions. Tension is a reflection of inability to continue or consummate behavior.

0. Extremely tense, strained, anxious, apprehensive at all times. Unable to relax; sleepless, restless, "on edge." Much muscular tension. Persistently heightened skeletal and visceral tension.

1. Usually quite anxious, frequently very tense. Has considerable difficulty relaxing. Experiences much strain, or lack of ease, when confronted with a new situation.

2. Feels mild tension much of the time. Becomes quite tense occasionally when faced with a new adjustment.

3. Usually free from unhealthy strain. Relaxes rather easily. Has reaction patterns which bring relief from tension when in new situations.

4. Seldom feels any tension except as a healthy stimulus to action. Feels almost no strain and is at ease very easily except in unusually challenging situations.

5. Feels almost no tension, strain, or anxiety. Perfectly able to relax at will.
10. **Empathy**

The term empathy includes the following concepts. Ability to feel with another person, or persons; more than just sensitivity. Vicarious reactions; having a large repertoire of reaction patterns in others that he can understand. Role-taking.

1. Absolutely no comprehension of the feelings of others, nor of their reactions.
2. Able to feel with a few persons in a very limited way.
3. Somewhat less than average ability to react vicariously, to take roles, and to understand the moods and actions of others.
4. A little better than average understanding of the feelings of others, and better than average ability to take roles.
5. Unusually good role-taking ability; can feel with many kinds of people in many emotions.
6. Nearly perfect understanding of the feelings and reactions of nearly all other people.

11. **Mood-outlook**

The term mood-outlook includes the following concepts. How one feels; hope, happiness, well-being; dejection, futility, gloom.

0. Completely hopeless and dejected. Always gloomy. Overwhelmed by the futility of life and effort.
1. Usually dejected, with little hope; occasionally mild optimism and weak cheerfulness.

2. A fair capacity for happiness, hope, and well-being, but easily and frequently discouraged.

3. Generally a somewhat more than average attitude of optimism, cheerfulness, and feeling of well-being.

4. Outlook is strong, confident, cheerful, and hopeful most of the time under most circumstances.

5. Very happy, cheerful, extremely optimistic; feels that anything is possible. Very great sense of well-being.

12. Independence

The term independence includes the following concepts. Ability to make up his own mind. Ability to make decisions and stick to them. Ability to stand up for his rights. Ability to stand alone in thought and action. (Not subject to bias or influence.) Determination to act; make up mind.

0. Completely dependent; can make no decisions and defends no personal opinions.

1. Generally dependent; makes a few simple decisions; seldom, and then only weakly, defends his rights or opinions. Usually afraid to have an opinion. Seldom trusts his
own judgment.
2. Makes many decisions easily on questions that are not too laden with emotion. Compromises his position more readily than the average person. Gives in rather easily if pressed.
3. In general, is a little better than average at making and holding to decisions, at defending his rights, and at thinking and acting independently.
5. Will stand alone against the world for his rights and principles; has very superior ability to make a decision and stick to it. Very rarely influenced by outside coercion.

13. Memory

The term memory includes the following concepts.
Ability to retain the present and recall the past. Freedom from preoccupation is involved. Mentally reproducing and recognizing past experiences.

0. Blank, no memory, no recall, or retention.
1. Retains and recalls a little; may be frequently preoccupied.
2. Little less than average ability to mentally reproduce and recognize past experiences; may be occasionally preoccupied.
3. A little better than average memory. Not often preoccupied.
5. Almost perfect retention and recall; no preoccupation.

14. **Spontaneity**

The term spontaneity includes the following concepts: Process of instigating and carrying on change and adjustment. Ability to initiate original action. Flexibility with a purpose. Ability to respond with interest. Reaction of the organism as a whole. Purposeful and conscious actions. Direction, orientation, and vital synthesis. It follows motivation. Must see a goal. Acting or proceeding from native feeling, without constraint or external force.

0. Initiates no original action; never acts of own volition. Inflexible; makes responses only as equal reactions to the stimulus.
1. Seldom acts from inner purposes; usually requires external influence. Very little flexibility.
2. Not quite average at initiating original action. Fairly good flexibility.
3. A little above average in ability to respond with interest, and to be flexible with a purpose.
4. Superior purposeful flexibility, adjustability, and ability to react with interest.
5. Extremely superior spontaneity.
RESULTS

The result of this research and this thesis is that an important new clinical rating scale has been developed which is both valid and reliable.

Table 1 shows that high item reliability was demonstrated for the Blackfoot Clinical Rating Scale, while table 2 shows that high total score reliability was also demonstrated. The reliability coefficients for total scores are .97, .97, and .96; for items they range from .76 to .98.

Total score validity was established by Larsen (17). He found correlations significant at the five per cent level between total scores on the Blackfoot Clinical Rating Scale and total scores on a modified Bales form. Item validity was obtained from the consensus of experts who agreed on the selection of the items, and who approved the wording of, and the concepts included in, the definitions.

Considerable support for the Blackfoot Clinical Rating Scale has come from members of the staff of State Hospital South in Blackfoot, Idaho. These staff members have given permission to have their opinions quoted here.

Dr. Martin, clinical psychologist, believes this rating scale "is definitely superior to any other rating scale of this sort now in existence."

Mr. Lake, clinical psychologist, believes "a compact rating scale of this kind is certainly needed in mental
Table 1. Coefficients of correlation of items on the Blackfoot Clinical Rating Scale. Coefficients were computed from the independent ratings by three judges (1) of twenty-six mentally ill persons.

<table>
<thead>
<tr>
<th>Item</th>
<th>A with B</th>
<th>A with C</th>
<th>B with C</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-esteem</td>
<td>.86</td>
<td>.89</td>
<td>.84</td>
<td>.86</td>
</tr>
<tr>
<td>Motivation</td>
<td>.94</td>
<td>.91</td>
<td>.92</td>
<td>.92</td>
</tr>
<tr>
<td>Self-confidence</td>
<td>.95</td>
<td>.96</td>
<td>.93</td>
<td>.95</td>
</tr>
<tr>
<td>Insight</td>
<td>.86</td>
<td>.82</td>
<td>.79</td>
<td>.82</td>
</tr>
<tr>
<td>Ability to abstract</td>
<td>.87</td>
<td>.87</td>
<td>.85</td>
<td>.86</td>
</tr>
<tr>
<td>Self-expression</td>
<td>.85</td>
<td>.89</td>
<td>.87</td>
<td>.87</td>
</tr>
<tr>
<td>Social interaction</td>
<td>.82</td>
<td>.83</td>
<td>.85</td>
<td>.83</td>
</tr>
<tr>
<td>Affective tolerance</td>
<td>.81</td>
<td>.80</td>
<td>.85</td>
<td>.82</td>
</tr>
<tr>
<td>Absence of tension</td>
<td>.96</td>
<td>.95</td>
<td>.98</td>
<td>.96</td>
</tr>
<tr>
<td>Empathy</td>
<td>.84</td>
<td>.79</td>
<td>.76</td>
<td>.80</td>
</tr>
<tr>
<td>Mood-outlook</td>
<td>.94</td>
<td>.93</td>
<td>.96</td>
<td>.94</td>
</tr>
<tr>
<td>Independence</td>
<td>.97</td>
<td>.95</td>
<td>.95</td>
<td>.96</td>
</tr>
<tr>
<td>Memory</td>
<td>.95</td>
<td>.94</td>
<td>.95</td>
<td>.95</td>
</tr>
<tr>
<td>Spontaneity</td>
<td>.86</td>
<td>.81</td>
<td>.84</td>
<td>.84</td>
</tr>
</tbody>
</table>

(1) A represents the clinical psychologist; B represents the psychiatrist; C represents the intern psychologist.
Table 2. Coefficients of correlation between the total scores of twenty-six patients rated independently by three judges (1) on the Blackfoot Clinical Rating Scale.

<table>
<thead>
<tr>
<th>Judges</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A with B</td>
<td>.97</td>
</tr>
<tr>
<td>A with C</td>
<td>.97</td>
</tr>
<tr>
<td>B with C</td>
<td>.96</td>
</tr>
</tbody>
</table>

(1) A represents the clinical psychologist; B represents the psychiatrist; C represents the intern psychologist.
institutions."

Dr. Prins, psychiatrist, said this scale "is a very good aid in reviewing a patient's condition and getting a more complete picture of his personality on a given day."

Dr. McCune, psychiatrist, voiced the opinion, "this scale is very important. It helps to bring to mind areas of the personality which might otherwise be easily overlooked."
CONCLUSIONS

A rating scale which is soundly constructed and thoroughly tested is a valuable tool for the clinician in a mental hospital.

From the study of the approved methods that were used in establishing the highly significant reliability and validity of the Blackfoot Clinical Rating Scale, this scale can be accepted as fulfilling the task for which it was intended. It is a valid and reliable instrument for evaluating and recording the personality development of patients with behavior disorders. In addition it serves as a frame of reference for diagnostic thinking and for planning changes in therapy when necessary.

In order to make the Blackfoot Clinical Rating Scale more efficient and to assure its continued usage, further research should probably be done to make the definitions more compact. The definitions should perhaps also be joined in some form directly to each rating scale sheet.
SUMMARY

1. The Blackfoot Clinical Rating Scale was prepared to meet the need for a convenient device for evaluating and recording personality changes in mentally ill patients.

2. After considerable deliberation by several trained clinicians, fourteen items were selected which included the most important areas of personality.

3. Rating levels from "0" to "5" were chosen as those permitting the best inter-item comparisons and providing sufficient scope and division to be effectively employed without being awkward.

4. A complete set of definitions was prepared which included all of the basic concepts under each item and clearly limited each rating to its particular level.

5. Very high reliability was established by computing the correlations between three judges, who independently rated the same twenty-six patients.

6. Validity, while not computed, is quite high. The high correlation between the ratings of the independent judges was possible only because the definitions were clear, valid descriptions of the personality areas which the judges were rating. The validity of the descriptions was assured by obtaining the consensus of expert clinicians who jointly approved the items, the wording, and the concepts.
7. Staff clinicians of State Hospital South in Blackfoot, Idaho were unanimous in their approval of the Blackfoot Clinical Rating Scale.

8. From this study it was concluded that the Blackfoot Clinical Rating Scale is a useful, valid, and reliable device for evaluating and recording personality changes in mentally ill patients.
BIBLIOGRAPHY


(17) Larsen, Reed M. *An Investigation Into the Relationship Between the Degree of Adjustment in Mentally Ill Patients and Three Experimental Social Situations.* Unpublished manuscript. Logan, Utah: Department of Psychology, Utah State Agricultural College, 1950.


