

A METHOD OF RATING FOREST SERVICE EMPLOYEES

by

ALVA WOODSON BLACKERBY

A THESIS

submitted to the

OREGON STATE COLLEGE

in partial fulfillment of  
the requirements for the  
degree of

MASTER OF FORESTRY

June 1939

APPROVED:

Redacted for privacy

Assistant Professor of Wood Products

In Charge of Major

Redacted for privacy

Chairman of School Graduate Committee

Redacted for privacy

Chairman of State College Graduate Council

### ACKNOWLEDGMENT

The writer wishes to express appreciation for the advice and criticisms offered by Mr. Robert Evenden of the School of Forestry. He is grateful to Mr. P. L. Paine, Assistant in the Division of Personnel Management of the United States Forest Service, for assistance and permission to conduct the testing experiment in the United States Forest Service. Mr. James McCarran, Secretary of the Multnomah County Civil Service Commission, kindly allowed the writer the use of equipment for scoring the Probst Service Report Rating Scales. Dr. R. J. Clinton of the School of Education contributed helpful suggestions. The writer also wishes to thank the other faculty members and students of the School of Forestry as well as other United States Forest Service employees who assisted.

## INTRODUCTION

### THE PROBLEM OF THE THESIS

The problem of this thesis is to develop a rating scale for professional foresters in the United States Forest Service from a survey of the general field of rating scales.

At present all of the regular Civil Service employees in the Forest Service, are rated once a year.<sup>1</sup> Although there are between 400 and 500 such employees in the North Pacific Region (Oregon and Washington),<sup>2</sup> the rating scale to be constructed will be designed to apply only to the professional foresters. This includes junior foresters, district rangers, foresters on the supervisors' staffs, and foresters on the staffs of the regional foresters.

Unfortunately human efficiency has to be measured. It can be most accurately measured in terms of physical production. For example, the best worker in a post hole dig-

- 
1. Two similar rating forms are used, Forms 3200 and 3201. Form 3201 is used for employees in the Professional and Scientific Service and grades nine to twelve, inclusive, of the Clerical, Administrative, and Fiscal Service. Form 3200 is used for the remainder of the employees which includes all employees in the Sub-professional Service, Clerical-Mechanical Service, Custodial Service, and grades one to eight, inclusive, of the Clerical, Administrative, and Fiscal Service. Sample copies of Forms 3200 and 3201 may be found in the appendix. See United States Civil Service Commission, Classification Statutes, Washington, U. S. Government Printing Office, Form 2910, June 1938, p. 8-15.
  2. Paine, P. L., Assistant in the Division of Personnel Management in the U. S. Forest Service--letter to the writer dated February 2, 1939.



ging gang is the one who digs the most post holes per day, other things being equal. But usually there are no production standards with which to rate the efficiency of foresters. The amount of service that a district ranger or a forest supervisor renders to the public through cooperation and education may be more important than the number of miles of trail built per month. In this case efficiency depends upon cooperation, initiative, leadership, etc., the qualities which are physically unmeasurable at present.

It is felt that the tools used for social measurement in the Forest Service are very much in need of improvement. Technicians in the physical science field measure distance to an accuracy of one millionth of an inch while measurements to one thousandth of an inch are commonplace. At the same time the personal or so-called unmeasurable qualities of employees are crudely and unscientifically measured. It is noteworthy when a supervisor, in the public service field, can measure the value of his subordinates with sufficient accuracy to place them in the groups to which they belong, when using a five class rating range.

Decisions relative to the maintenance of tools in a modern repair shop are more scientifically made than decisions vitally affecting human life and happiness--promotions, demotions, transfers, etc.

After examining the rating scale field it is planned

in this thesis to develop a rating scale that will be particularly applicable to the measuring of the service value or total efficiency of foresters in the United States Forest Service.

## TABLE OF CONTENTS

	Page
INTRODUCTION . . . . .	1
PART I	
INTRODUCTORY SURVEY OF THE RATING SCALE FIELD	
I. STATEMENT OF THE RATING PROBLEM . . . . .	1
II. PURPOSES AND USES OF RATING SCALES . . . . .	6
Aiding Management . . . . .	6
Educating the Raters . . . . .	11
Stimulating and Educating the Ratee . . . . .	13
Advertising Personnel Management . . . . .	15
Furnishing Research Material . . . . .	16
III. SPECIAL ADVANTAGES OF RATING SCALES . . . . .	17
Requires Analytical Judgment . . . . .	17
Makes Standards Uniform . . . . .	18
Makes Ratings Definite . . . . .	18
Makes Ratings Understandable and Comparable . . . . .	19
Standardizes Ratings . . . . .	19
Systematizes Ratings . . . . .	20
Simplifies the Rating Job . . . . .	21
IV. DIFFICULTIES INHERENT IN RATING SCALES . . . . .	22
Halo . . . . .	22
Poor Sampling . . . . .	23
Superficial Judgments . . . . .	24
Favoritism . . . . .	25

# TABLE OF CONTENTS (Continued)

	Page
High and Low Raters . . . . .	25
Discrimination Overloaded . . . . .	26
Length of Acquaintance . . . . .	27
Poor Rater Instructions . . . . .	28
Pseudo-Sciences . . . . .	28
V. RATING SCALE CRITERIA . . . . .	30
Objectivity . . . . .	30
Reliability . . . . .	32
Validity . . . . .	34
VI. REACTIONS TO RATING SCALES . . . . .	38
VII. EARLY RATING METHODS AND DEVELOPMENT . . .	43
VIII. TYPES OF RATING SCALES . . . . .	46
Check-List Rating Scale . . . . .	46
Numerical Rating Scale . . . . .	47
Alphabetical Rating Scale . . . . .	48
Adjective Rating Scale . . . . .	49
Descriptive Rating Scale . . . . .	50
Multiple-Step Rating Scale . . . . .	51
Linear Rating Scale . . . . .	52
Order of Merit Rating Scale . . . . .	53
Scott Man-to-Man Comparison Scale . . .	55
Graphic Rating Scale . . . . .	57
Probst Service Report Rating Scale . . .	61

## TABLE OF CONTENTS (Continued)

	Page
Service Rating Forms Used in the United States Forest Service . . . . .	66
Horizontal and Vertical Ratings . . . . .	73
PART II	
DEVELOPMENT OF A RATING SCALE	
IX. NEED OF A SPECIAL RATING SCALE FOR EACH GROUP . . . . .	74
X. SELECTING THE ESSENTIAL TRAITS OF A SUCCESSFUL FORESTER . . . . .	76
XI. TABULATION OF THE TRAITS . . . . .	92
XII. DEVISING THE RATING SCALE . . . . .	99
Features of Sound Rating Scales . . . . .	99
Approved Administrative Features . . . . .	109
Weighing the Essential Qualities . . . . .	115
Defining the Essential Qualities in Terms of Observable Behavior . . . . .	118
Developing a Method of Recording Comparative Degrees of Behavior . . . . .	120
Direction to the Rater . . . . .	123
Scoring Methods . . . . .	124
Converting Total Numerical Scores Into Final Ratings of Five Classes . . . . .	125
Providing an Opportunity for Special Remarks . . . . .	129
Final Rating Scales . . . . .	130
XIII. TESTING THE RATING SCALE . . . . .	145

## TABLE OF CONTENTS (Continued)

	Page
Comparing the Service Record and the Man-to-Man Record Rating Scales . . . . .	145
Rating Scale Testing in the School of Forestry . . . . .	147
Testing the Rating Scales in the Forest Service . . . . .	150
XIV. SUMMARY . . . . .	166
Recommended Subjects for Additional Study	172

### TABLES, FIGURES, AND FORMS

Table I.	Evaluation by Rank Order of Six Essential Qualities in Foresters .	115
Table II.	Distribution of Numerical Scores Made by Forest Supervisors and Assistant Forest Supervisors in Rating Their Professional Personnel . . . . .	126
Table III.	Conversion Table for Changing Total Numerical Scores to Class Scores .	129
Table IV.	Ratings Made by Nine Students and One Professor in the School of Forestry of One Man with the Service Record and Man-to-Man Record Rating Scales	146
Table V.	Ratings Made in January by Four School of Forestry Professors of Students, with Three Rating Scales . . . . .	178
Table VI.	Ratings Made in March by Four School of Forestry Professors of Students, with Three Rating Scales . . . . .	179
Table VII.	Scores Made by Forest Service Raters Using Guide For Using Form 3200, Probst Service Report, and Service Record Rating Scales . . . . .	180

## TABLE OF CONTENTS (Continued)

	Page
Table VIII. Class Ratings Converted From Numerical Ratings Shown in Table VII . . . .	181
Table IX. Total Numerical Scores and Average Score Per Item Checked, Made by Forest Service Raters on the Odd and Even Numbered Questions Using Guide For Using Form 3200 Rating Scale .	182
Table X. Total Actual Numerical Scores and Average Score Per Item Checked, Made by Forest Service Raters on Odd and Even Numbered Questions Using Service Record Rating Scale . . . . .	184
Table XI. Conversion Table for Changing Total Numerical Scores into Final Ratings	186
Figure I. Letter Requesting Traits and Objectives for Rating Scales . . . . .	76
Figure II. Traits and Behaviors Questionnaire .	77
Figure III. Distribution of Forty-two Ratings on Twenty-five Men in the Forest Service	154
Rating Scales:	
Service Record . . . . .	131
Man-to-Man Record . . . . .	138
Service Rating Form 3200 . . . . .	187
Service Rating Form 3201 . . . . .	188
Reverse Side of Forms 3200 and 3201 . .	189
Guide For Using Service Rating Form 3200	190
Probst Service Report . . . . .	195
Reverse Side of Probst Service Report .	196

PART I

INTRODUCTORY SURVEY OF THE RATING SCALE FIELD



## CHAPTER I

### STATEMENT OF THE RATING PROBLEM

# A METHOD OF RATING FOREST SERVICE EMPLOYEES

## CHAPTER I

### STATEMENT OF THE RATING PROBLEM

After having employed an individual for a given job, the problem of evaluating this person's worth to the organization soon arises. Is his production satisfactory? Is he promotional material? How does he get along with his fellow employees? These questions, and others similar in nature, must be answered by those responsible for the supervision of personnel activities in any organization.<sup>3</sup>

Rating scales are designed to measure reactions of human beings to human beings when employee value cannot be based on production.<sup>4</sup> Measured production is inapplicable in many cases and in most government positions.<sup>5</sup> However, there are many human reactions and characteristics of employees, the measurement of which seems important to management, for which there are no objective measures.<sup>6</sup> Certain types of human qualities, of which initiative, persistence, leadership, self-control, perspective, and

---

3. Clarke, Walter V., "Rating Employees", *Journal of Personnel Research*, V. 15, p. 100.

4. Laird, Donald A., The Psychology of Selecting Men, p. 100.

5. Feldman, Herman, A Personnel Program for the Federal Civil Service, A report by the Personnel Classification Board, 1931.

6. Yoder, Dale, Personnel and Labor Relations, p. 249.

appearance are fairly representative, may be quite important to successful performance in particular positions. Such traits cannot be stated in terms of items per minute, or some other quantitative unit, but they are nevertheless important in vocational prognosis.<sup>7</sup>

Employees doing the same class of work differ in these human qualities as well as in day-by-day performance<sup>8</sup>, and these qualities have an important bearing on the evaluation of services. Management is therefore interested in their presence or absence but "... there are few if any objective standards ... available to provide a quantitative measurement of these characteristics. It is necessary, therefore, to rely on judgments and opinions"<sup>9</sup>.

"Opinion is frequently faulty. In more cases than otherwise it is formed according to interest and prejudice."<sup>10</sup> In daily contacts between supervisors and subordinates, personality, consciously or unconsciously, affects the supervisor's opinion of each employee. We are all apt to be more generous and lenient toward a person for whom we entertain high regard or affection. Therefore, means need to be made available whereby those opinions may

7. Burt, Harold Ernest, Principles of Employment Psychology, p. 317.

8. Baridon, Felix E., and Loomis, Earl H., Personnel Problems, p. 185

9. Yoder, Dale, op. cit., p. 249.

10. Scott, Walter Dill, Clothier, Robert C., and Mathewson, Stanley B., Personnel Management, p. 101.

be formed as fairly as possible. Without such means, snap judgment is apt to take the place of deliberate appraisal. Personal likes and dislikes are apt, unconsciously, to lead to erroneous conclusions.

Burt says that men are constantly observing one another and from external behavior, inferring something regarding mental traits. These opinions are of somewhat dubious value, especially in the form in which they are most frequently available--usually some glittering generalities to the effect that he is a "good man" or a "poor man", or "does not take hold". Such terms are quite relative and mean radically different things to different persons. Being a good man in the estimation of one person may be equivalent to mediocrity in the estimation of another. Mr. Burt further explains that general impressions of this sort are likewise apt to reflect prejudice. If the rater has had some unfortunate experience with the individual in question--for example, if he has encountered some single instance of carelessness, he is apt to ascribe the bad impression of this incident to the individual's entire personality. Hence it is desirable to abstract somewhat from these prejudices and general impressions and obtain the estimates in more scientific fashion.<sup>11</sup>

The U. S. Forest Service operates, theoretically at

---

11. Burt, Harold Ernest, op. cit., p. 366.

least, on the principle of merit and if a merit system is to operate within the service as well as at its portals, account must be taken of the efficiency of an employee's work.<sup>12</sup> But it must be recognized that if we are to have a merit system, the fact of merit must be ascertained by some means. If no dependable aids are available for this purpose, the decision will rest upon the subjective and often unreliable estimate of a single superior. Such estimates and judgments are notoriously unreliable. As pointed out by Samuel H. Ordway, Jr., and John C. Laffan, "Evidence of the danger of error inherent in judgment rating is universal and overwhelming".<sup>13</sup>

Some method is needed that will afford a more uniform method of expressing opinion regarding employees that deals less with general impressions or prejudice and more with specific traits. This can be accomplished to a certain extent by rating the traits separately and combining them into a final rating.

Rating scales or service ratings<sup>14</sup> have been variously described as "... merely a device for keeping all of the essential qualities to be considered before the rater, which tends to produce more uniform, consistent, and compa-

---

12. Mosher, William E. and Kingsley, J. Donald, Public Personnel Administration, p. 427.

13. Ordway, Samuel H. Jr., and Laffan, John C., "Approaches to the Measurement and Reward of Effective Work of Individual Government Employees", National Municipal Review, supplement, V. 24, 1935, p. 559.

14. See bottom page 5.

nable judgments".<sup>15</sup> It yields a record of the rater's subjective estimate of the person's ability or traits.<sup>16</sup> Or it might be thought of as a systematized record of opinion. A purpose is inferred when it is said that the rating scale is a method of recording estimates of abilities in a quantitative and comparable way.<sup>17</sup>

Mr. Arthur W. Kornhauser says, "Rating scales are simply convenient forms for securing more adequate personal estimates of people that are obtained by less formal methods. The ratings are typically estimates given periodically by executives, supervisors, or teachers concerning the workers or students under them."<sup>18</sup>

14. According to J. B. Probst in his book, Service Ratings, p. 10, the term "service rating" is now preferred by personnel administrators to its predecessor, "efficiency rating". The new term is more expressive and broader in its application. "A service rating", says Mr. Probst, "properly embraces the factors of efficiency, character, and conduct and attempts to indicate by means of a symbol, word, letter, or per cent, the measurement of an employee's value in relation to his fellow workers."
15. Cleeton, Glen U., and Mason, C. W., Executive Ability, Its Discovery and Development, p. 51.
16. Bingham, Walter Van Dyke, and Freyd, Max, Procedures in Employment Psychology, p. 87.
17. Ibid., p. 122.
18. Kornhauser, Arthur W., "What Are Rating Scales Good For?", Journal of Personnel Research, Vol. V, p. 189-193.

## CHAPTER II

### PURPOSES AND USES OF RATING SCALES

## CHAPTER II

### PURPOSES AND USES OF RATING SCALES

The most obvious purpose of any rating system according to Kornhauser is the securing of personal information that will be useful to administrative officers engaged in personnel activities.<sup>1</sup> In addition, however, important results are achieved by the rating process in the effects it has on the people rated and on those doing the rating. The educational and inspirational effects of rating systems may, indeed, be as important as the obtaining of information about the individuals judged. Tead says that this is, perhaps, its most valuable use.<sup>2</sup>

It is certainly true that these less obvious results deserve careful consideration. The procedure may also have a distinct value in disseminating the personnel viewpoint and in supplying data for personnel research. These several points are briefly sketched, as follows:

#### Aiding Management.

Service ratings are an aid to management and administrative officers. The ratings furnish useful information about characteristics and abilities of employees. Attention is called to individuals who are exceptionally strong or exceptionally weak in certain traits which, on the one

---

1. Kornhauser, Arthur W., op. cit., p. 189-193.

2. Tead, Ordway, Human Nature and Management, p. 217.



hand, may indicate cases of maladjustment where transfer or special training is requisite, and, on the other hand, serve to locate promotional material. Also, attention is called to cases where the person is rated markedly high or low by particular raters. Furthermore, successive ratings lead to a recognition of marked improvement or retrogression. This information makes possible better judgments in administrative decisions concerning the individual in specific instances such as:

1. To differentiate between who is and who is not to be promoted, or if conditions make such action necessary, those who are to be demoted. Some method of periodically reporting upon the work of the employee is essential to any sound promotion plan.<sup>3</sup> Mosher and Kingsley point out that there is probably no more crucial test of the success of the personnel program than the handling of promotions.<sup>4</sup> With the possible exception of policy determining positions the sole basis of promotion must be fitness. In the interest of service efficiency, no other policy is defensible. As one proceeds from the lower to the higher supervisory and executive positions, service ratings and records become increasingly important as a basis for promotion because the only real test for administrative

---

3. Mosher, William E., and Kingsley, J. Donald, Public Personnel Administration, p. 268.

4. Ibid., p. 265.

capacity is performance in office, and the only method of estimating it is the personal opinion of those who know the candidate. Mosher and Kingsley add, "But this does not mean that all attempts at objectivity and the elimination of personal bias should be given up. Rather, it means sustained efforts to perfect a record system and a service rating scheme, which will reduce the influence of irrelevant and undesirable factors to a minimum.<sup>5</sup> However, even the most fool-proof method of rating the efficiency of the employee in his present job is entirely inadequate<sup>6</sup> as the sole criterion for promotion to a position involving increased responsibilities. Past efficiency must be considered as but one factor among many.

2. To provide one basis for making salary increases within the scale of pay for a given class.<sup>7</sup>

3. To determine the order of layoff and reemployment.<sup>8</sup>

4. To ascertain if a worker's value is increasing or decreasing.<sup>9</sup>

5. To discover employees who ought to be transferred by aiding supervisors in distinguishing strong and weak points in individuals.

---

5. Ibid., p. 274.

6. Ibid.

7. Probst, J. B., Service Ratings, p. 10.

8. Ibid.

9. Ibid.

6. To help the supervisors to determine the type of training or development needed to increase the worth of individual employees. It acts as an educational weapon in enabling the personnel executive to discuss with the person rated his particular deficiencies and the ways of overcoming them.<sup>10</sup>

7. To grant special privileges.

8. In making wage adjustments. A basis is provided for agreement or intelligent disagreement.<sup>11</sup>

9. For use in selection. The important elements in a particular position are emphasized, thus calling attention to the characteristics needed in new employees.

Before leaving the discussion on how management is served by service ratings it might be well to review the attitude and actions of the U. S. Forest Service along this line. In answer to a question involving the use of rating scales in relation to salary increases, salary reductions, demotions, and dismissals, a Forest Service officer stated, "We have a scale, based on individual efficiency ratings, which determines whether or not an employee is entitled to promotion, demotion, or dismissal". The regulations,<sup>12</sup> specifically, are as follows:

---

10. Tead, Ordway, op. cit., p. 217.

11. Watkins, Gordon S., Labor Management, p. 328.

12. United States Civil Service Commission, Preparation of Efficiency Ratings, p. 9-10.

"Rules for salary increases. The following rules are prescribed to establish ratings which employees must attain to be eligible for salary increases:

- (1) "An employee attaining a final rating<sup>13</sup> of 'Excellent' or 'Very Good' is eligible for a salary increase within the grade to which his position is allocated if he is not already receiving the maximum pay rate of his grade.
- (2) "An employee attaining a final rating of 'Good' is eligible for a salary increase up to, but not beyond, the middle salary rate<sup>14</sup> of the grade to which his position is allocated, but if he is already receiving a salary higher than the middle rate of his grade he is not subject to a salary reduction on that account.

"Rules for salary reductions, demotions, or dismissals. The following rules are prescribed to establish ratings which no employee may receive without being demoted to other duties or reduced in salary, or both, or dismissed for inefficiency:

- (1) "An employee whose final rating is 'Unsatisfactory' will not be continued longer in the work upon which he was engaged during the period covered by the rating. He will be assigned to duties more nearly commensurate with his ability, in a lower classification grade, and his compensation will be fixed at a rate not in excess of the middle salary rate for such grade. Should no suitable vacancy be available in a lower classification grade, he will be separated from the service for inefficiency.
- (2) "An employee whose final rating is

---

13. There are five possible final ratings: Excellent, very good, good, fair, and unsatisfactory.

14. Each grade has seven salary rates. The fourth salary rate will be considered the middle salary.

'Fair' and who is receiving a salary above the middle salary rate of the grade to which his position is allocated will be demoted one salary step, but if he is already receiving a salary equal to or below such middle salary rate, he is not subject to a salary reduction on that account."

A Forest Officer was asked; "After a service rating indicates employee weaknesses, is any attempt made to aid the employee in remedying these weaknesses or in making some adjustment?".

His reply was, "Yes. When an employee begins to fall down, his immediate superior goes over the case with the employee in an endeavor to determine what is wrong, and devise methods of overcoming the situation. Sometimes this means assignment of an employee to other lines of work. If the employee is a misfit, the only course is removal."

#### Educating the Raters.

The making of these personal judgments has a beneficial effect on the persons doing the rating. It impresses the executives and supervisors with the importance of thinking of their employees not in general terms, but in terms of the particular qualities which management regards as important. It impresses them with the inadequacy of judgments which are vague, not specific. The procedure brings prominently to the attention of supervisors and executives, the importance of knowing their workers as

individuals. It teaches executives to think of their employees in a deliberate analytical manner; it leads them to avoid snap judgment.<sup>15</sup> It serves to prevent some isolated instance of good or bad performance from coloring unduly the executive's judgment of the employee's general worth. The necessity of giving estimates or judgments on personal characteristics leads the rater to a thoughtful analysis of the people rated and tends to make him more alive to his opportunities and responsibilities in developing men.<sup>16</sup>

Burt shows the educational value to the rater in the following words:<sup>17</sup>

Management "... observes the man rated more closely if he is required occasionally to rate him. In addition to arousing personal interest in the man it leads the rater to observe him with reference to different traits and consider them separately. The natural tendency is to devote attention primarily to the man as a whole or to some outstanding aspect. It is easy to dislike a man's face and overlook his other good qualities. The rating scale calls attention to these other qualities and teaches one to observe them too. One may discover that after all the man is rather skillful, ingenious, and cooperative. On the other hand, the scale may call attention to the man's laziness which had been previously overshadowed by his affability. In this way one's final opinion of the man and one's whole attitude toward him may be very appreciably changed. Furthermore, this procedure keeps the whole notion of personality alive in the mind of the executive."

15. Scott, Walter Dill; Clothier, Robert C., and Mathewson, Stanley B., op. cit., p. 102.

16. Kornhauser, Arthur W., op. cit., p. 191.

17. Burt, Harold Ernest, op. cit., p. 319.

### Stimulating and Educating the Ratee.

The use of service records and ratings has both a stimulating and an educational effect on the persons who are rated. The most important of all non-financial incentives is the opportunity for growth and the stimulus to grow. For a young person this often outranks any other consideration.<sup>18</sup> This takes two forms:

1. If the individual knows that his abilities and characteristics are being studied and recorded he is apt to strive to make a good impression. This may encourage a certain amount of self-analysis and evaluation and he may seek to determine his weak points with a view to improvement.<sup>19</sup> He may also realize that the ratings have something to do with his status in the concern. This will serve as additional motivation and be an incentive to do as effective work as he can.

2. When and if the estimates are made known to the ratee<sup>20</sup> he can gain some useful views of himself. A worker may be a gross offender with respect to cooperativeness or

---

18. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 253.

19. Burt, Harold Ernest, op. cit., p. 319.

20. In the Forest Service "the head of each department or other proper administrative authority shall, upon completion of Efficiency Rating Form No. 4B, notify each employee under his jurisdiction of his final efficiency rating in terms of the appropriate adjective". The adjective referred to is one of the five--excellent, very good, good, fair, or unsatisfactory.<sup>21</sup>

21. U. S. Civil Service Commission, op. cit., p. 9.

dependability and yet never be aware of the fact. There may be distinct reasons for his failure to progress, yet his reaction has only been that he is not getting ahead but he doesn't know why. He may think that the boss "has it in for him". The material is likely to make him do some real thinking about himself and his possibilities. This self-analysis in some cases will lead to serious consideration of vocational aptitudes and plans.<sup>22</sup> It may also suggest the problem of developing a more effective personality and lead to efforts toward improvement in that direction. Scott, Clothier, and Mathewson further explain this last point as follows:<sup>23</sup>

"... it impels the employee to exercise conscious effort in improving himself in those specific qualities deemed important by the management. The fog and uncertainty which beclouds the mental processes of workers with respect to their status in their companies is often not apparent to the management. The members of the management have so grown away from the viewpoint and habit of thought of the workers down the line that they have forgotten the discouragements, the gropings, the sense of futility which so often assail them. This kind of discouragement is largely the product of uncertainty as to what the management expects of them and as to the degree to which they are fulfilling such expectations.

"It is refreshing to observe the reassurance with which employees react to such a plan when it is properly installed. For the first time they see in black and white certain specific qualities the management wants

---

22. Kornhauser, Arthur W., op. cit., p. 191.

23. Scott, Walter Dill, Clothier, Robert C., and Mathewson, Stanley B., op. cit., p. 102-3.



them to possess. Direction is given to their efforts for self-development. Their real objectives appear out of the mist."

It discourages bluffing. "... They observe for the first time a practical instrument whereby their strengths and failings can be known to the management..... Here we have a distinct incentive. The employees know what is expected of them. They know that management is cognizant of the extent to which they do or do not possess such qualities. They are informed that the appraisals are periodic and they are stimulated to improve their records by conscious self-cultivation."

#### Advertising Personnel Management.

The rating activities help to introduce, reenforce, and keep alive, the personnel spirit. Specific personal devices and records may become somewhat distasteful to some executives. It is therefore important that a persuasive personnel spirit and an appreciation of personnel objectives permeate the organization. The use of rating scales helps significantly in these large aims because the rating scale concerns everyone--those rated, those rating, and those using the rating. It provides a natural and useful link between personnel officers and other members of the organization, leading readily to discussions of the use of rating scales in dealing with people individually and personally--thence to the whole personnel philosophy. It is interesting to note that the personnel system entered the Army in 1919<sup>24</sup> by the rating scale route.

---

24. Kornhauser, Arthur W., op. cit., p. 191.

Furnishing Research Material.

Rating results furnish useful material for personnel research. Problems of selection and turnover continually call for study of the personal qualities which determine success or failure. Rating scales aid in ascertaining the personal traits which differentiate good workers from poor workers<sup>25</sup> and these can be used in checking the results of selective tests and devices. Kornhauser adds, "Ratings may likewise be valuable as a technique for studying existence of 'personality types', the interrelatedness of traits, the influence of various factors in determining our estimates of people, and so on".<sup>26</sup>

---

25. Ibid.

26. Ibid.

### **CHAPTER III**

#### **SPECIAL ADVANTAGES OF RATING SCALES**

### CHAPTER III

#### SPECIAL ADVANTAGES OF RATING SCALES

The effects of rating systems which have been discussed are true in some degree of any procedure for obtaining estimates of individuals' characteristics. However, the standardized rating scale is ordinarily believed to possess certain special merits as compared with informal and unstandardized personal estimates. Some of the more important points follow:

Requires Analytical Judgment.

The rating scale is analytic; it calls for judgment on a variety of distinct and defined character traits.<sup>1</sup> These traits have been carefully chosen as the essential ones for the purpose. The analytic nature of the scale prevents the rater's seeing only one or two striking characteristics and omitting others. It makes it easier for the executive to think of his subordinates in specific terms. By these methods the executive appraises him in one quality alone without, at the moment, considering him in other qualities. He must stop and think instead of using merely his first off-hand opinion. It also tends to keep him from letting a single general impression bring him to specific strong and weak points in the person rated.

---

1. Kornhauser, Arthur W., op. cit., p. 192.

This may be summarized as resulting in:

1. Requiring the supervisors to give employees more careful consideration, thus minimizing the effects of personal likes and dislikes.<sup>2</sup>
2. Helping the supervisors to distinguish strong and weak points in individuals so that better adjustments of the personnel can be made in each organization.<sup>3</sup>
3. Giving a more detailed picture of the worth of an employee as compared to others or as compared with former ratings.<sup>4</sup>

Makes Standards Uniform.

The standards by which executives and supervisors judge their workers are made uniform. One executive does not emphasize one set of qualities and another an entirely different set of qualities as the personal qualities the company or concern considers important are set forth.<sup>5</sup>

Makes Ratings Definite.

The ratings are explicit, unambiguous, and they are recorded.<sup>6</sup> This means that the estimates can be studied, compared, and checked over. The ratings can be clearly interpreted; they are not vague statements which can mean

---

2. Baridon, Felix E., and Loomis, Earl H., op. cit., p. 187.

3. Ibid., p. 187.

4. Ibid.

5. Scott, Walter Dill, Clothier, Robert C., and Mathewson, Stanley B., p. 179.

6. Kornhauser, Arthur W., op. cit., p. 192.

almost anything that one pleases to read into them. Also, the rater is likely to be more careful when he recognizes that inaccurate ratings will reflect upon him.

Standardizes Ratings.

The records are uniform and standardized. Estimates given by different raters and estimates made at different times are directly comparable. These records can be readily compiled and summarized for use.

Makes Ratings Understandable and Comparable.

The records of ratings are quantitative.<sup>7</sup> This makes especially feasible the recording, combination, and statistical treatment of the data. Judgments of different raters are easily compared with one another, averaged, or checked against independent facts concerning the person rated. Bad rating tendencies can be detected and, in some measure, corrected. It is possible to apply compensations, when needed, for an executive's inherent tendency to think generously of his people on the one hand or to think of them with exacting severity on the other.<sup>8</sup>

It becomes possible to find out quantitatively how reliable and usable the personal ratings are. Arthur W. Kornhauser, of the University of Chicago, adds:<sup>9</sup>

---

7. Ibid.

8. Scott, Walter Dill, Clothier, Robert C., and Mathewson, Stanley B., op. cit., p. 101-2.

9. Kornhauser, Arthur W., op. cit., p. 192.

"Moreover, those rated can be directly compared with one another and the relative position of each can be stated. This is to be contrasted with the results from a non-quantitative plan of estimates. With the rating scale, such statements as this are made possible: 'In intelligence, this man is in the top ten per cent of his group, according to the average rating of three superiors'. Without the quantitative scale, we would be able to say: 'One judge states that the man is "exceptionally alert and keen". Another says he is a "brilliant fellow", etc.' The trouble is that we do not know how these various phrases are related to one another."

#### Systematizes Ratings.

Ratings are made systematically and periodically.<sup>10</sup> They conduce to deliberate rather than hasty judgments. Provision is made for definitely recording a series of opinions which may be used as a check against snap judgment made under particularly irritating circumstances.<sup>11</sup> Unless the employees are few in number, the administrative officer without such aids is apt to overlook or forget past incidents, facts, or pieces of work, which may have an important bearing upon the determination of relative merits of the several employees under consideration for promotion or other adjustments.<sup>12</sup> In other words, it helps the executive base his judgment of subordinates not upon some recent outstanding instance of good or poor per-

---

10. Ibid.

11. Watkins, Gordon S., op. cit., p. 328.

12. Mayers, Lewis, The Federal Service, p. 304.

formance, but upon the character of his performance day-in and day-out.<sup>13</sup>

Where this is not the case, judgments are likely to be too much colored by some recent striking achievement or failure on the part of the person rated.

It periodically brings to the attention of management the qualities possessed by individual employees and the degree in which each is improving his standing in those qualities. It makes it possible for those whose development is rapid or whose ability is unusual to be more readily considered by management for a wage increase, promotion, or assignment to work in which his superior qualities can be utilized.<sup>14</sup>

Simplifies the Rating Job.

Some rating scales are extremely easy to use.<sup>15</sup> At least the rater does not have to rack his brain for adequate adjectives with which to describe a ratee nor to hunt for something to say.

---

13. Scott, Walter Dill, Clothier, Robert C., and Mathewson, Stanley B., op. cit., p. 178-9.

14. Ibid.

15. Kornhauser, Arthur W., op. cit., p. 192.



## CHAPTER IV

### DIFFICULTIES INHERENT IN RATING SCALES

#### CHAPTER IV

#### DIFFICULTIES INHERENT IN RATING SCALES

As has been pointed out before, we are constantly, consciously or unconsciously, judging each other. Bias and lack of uniformity in judging employees throughout an organization often creeps into raters' judgments. Some of the more important difficulties in the way of executives judging their subordinates with justice will now be considered.

#### Halo.

The fundamental weakness of service ratings, according to Leonard D. White, Professor of Political Science in the University of Chicago, is that too little is known of the human traits which are conducive to efficiency in given lines of work. As a result there has been a resort to vague and comprehensive terms such as character and cooperativeness which in effect mean very little when put at work in concrete situations. It is difficult to set a measurement upon the degree of character a person possesses, or the degree of cooperativeness which he demonstrated with others. It is difficult to think of a person in terms of one specific quality without letting our opinion of him in other qualities color our judgment of him in that quality.<sup>1</sup>

---

1. See bottom page 23.

For example, if an employee is obviously industrious and turns out great quantities of work, we may unconsciously be led to think of him as possessing initiative in high degree when a careful analysis of the facts may reveal that he is entirely lacking in this important quality.

Thorndike has discussed this tendency under the title of "halo". Yoder stresses it by saying, "This is by all odds the most common deficiency in ratings". He adds that an individual appears superior, inferior, or average, in a general way, and all individual qualities tend to be rated at that level. If the rater regards the rates as a "poor man", he rates him low on all counts.<sup>3</sup> Or, an extraordinarily high rating on some one trait may shed such a luster over the other qualities as to raise the rating on them to an unwarranted degree.

#### Poor Sampling.

The sampling is often poor. General opinions are in many cases based upon an isolated instance in an employee's experience. Any spectacular instance of good or poor performance often prejudices the superior in favor of the employee or against him, as the case may be, to a disproportionate degree.<sup>4</sup>

---

1. White, Leonard D., Introduction to the Study of Public Administration, p. 495.

2.

3. Yoder, Dale, op. cit., p. 274.

4. Scott, Walter Dill, Clothier, Robert C., and Mathewson, Stanley B., op. cit., p. 175.

A minor error may, in a particular instance, prove especially irritating and costly, regardless of the fact that the worker who committed the error may not have committed a similar error in a "blue moon" and probably will not commit another in an equally indefinite period of time, but it is probable that the single error will cost him dearly in reputation for accuracy in the judgment of his superior. Again, an employee who never had an idea before in his life may stumble upon a suggestion which finds favor with the boss; it is probable that this isolated incident will boost his stock in his superior's judgment far more than the everyday facts warrant.

#### Superficial Judgments.

Insufficient time and study may be devoted to the rating process. The average executive is working under pressure. He has to get his work out. He has no time, so he thinks, to spend in refinement of judgment. He is inclined to decide that a worker is good or no good by production or promptness with which reports are submitted alone.<sup>5</sup> He is disinclined to take the time to determine why the worker is making good or why he is failing to do so. He fails to get acquainted with the men he is rating or fails to analyze and report carefully and accurately what he does know about him.<sup>6</sup>

---

5. Ibid.

6. White, Leonard D., op. cit., p. 316.

### Favoritism.

Everyone is influenced unconsciously by his likes and dislikes. To get the rater to take an impartial, objective attitude toward the task and toward employees is a difficult undertaking.<sup>7</sup> Biased ratings, due to conscious prejudice or partiality, are probably far less frequent than unintentional bias, traceable to some amiable quality in the employee, good appearance, tact, etc., or to a brusque manner, unpleasant voice, or other socially irritating trait.

As a rule, so long as the behavior of the ratee, in its many aspects such as mannerisms, clothes, work, habits, etc., is in line with or agrees with the pleasant experience of the rater, a favorable rating may be expected. But when a ratee's behavior reminds the rater of unpleasant experiences, either remembered or forgotten, the rating is apt to be low.<sup>8</sup>

### High and Low Raters.

Some raters are severe while others are lenient. One executive does not think in the same terms as another. One may be generous and liberal in his judgments; another may be severe and exacting. Even if they think alike, one may

---

7. Kingsbury, Forrest A., "Analyzing Ratings and Training Raters", Journal of Personnel Research, V. 1, p. 377-383.

8. Dashiell, John Frederick, Fundamentals of Objective Psychology, p. 179.

express his judgment in lenient terms; and the other may express his less generously.

Again, one may lay special emphasis on one quality, such as industry; another may lay special importance upon another such as cooperativeness; the third may lay special importance upon another, such as neatness.

Scott, Clothier, and Mathewson say, "Obviously, it is difficult without the proper instruments to insure that the employees throughout an organization, working under many different executives, will be judged uniformly. Because this is true, an element of injustice enters which tends to defeat the purposes of management and is unfair to the employees and to the executives who are being judged."<sup>9</sup>

#### Discrimination Overloaded.

The power of discrimination is often overloaded. Raters cannot be expected to distinguish between more than five degrees of quality.<sup>10</sup> Often this is complicated by expecting raters to judge a ratee on several traits at once. For example a rater may be expected to judge the quality of initiative. At first glance it may appear that the rater is evaluating only one trait when in reality initiative includes intelligence, dependability, and

---

10. Yoder, Dale, op. cit., p. 266.

energy,<sup>11</sup> making the task a very difficult undertaking. Probst says that one of the principal defects inherent in the rating and reporting systems of the past is in requiring the rater to analyze, compare, and evaluate a number of individual qualities at the same time.<sup>12</sup>

#### Length of Acquaintance.

Judgments are affected by the length of acquaintance. If the length of acquaintance has been long the rater is apt to give too favorable an estimate due to unconscious identification of the older subordinates with himself, hesitation to concede that his long influence has not improved them, and adaptation to their weak points.<sup>13</sup> Laird says, "... the longer acquainted one is with the person he is recommending the poorer his estimates are". On the other hand he points out that the raters should be well acquainted with those they are rating.<sup>14</sup>

Also, first impressions based upon general impressions are notoriously unreliable. When raters judge employees or prospective employees on the basis of first impressions there is little agreement between the opinions of different raters or between the initial opinions and judgments made after a reasonable period of acquaintance. This unrelia-

---

11. Meyers, Harry, Human Engineering, p. 203.

12. Probst, J. B., op. cit., p. 16.

13. Laird, Donald A., op. cit., p. 154.

14. Ibid., p. 193.

bility of first impressions is largely a result of associations with peculiarities of gait, posture, clothes, and expression which may develop a tendency to react favorably or unfavorably to these stimuli. Care must be taken to see that inaccurate initial opinions are not allowed to become permanent.<sup>15</sup>

#### Poor Rater Instructions.

A rater may lack proper training and experience. Often the person making the rating scale forgets that few executives have had training and practice in analyzing and thinking out specific aspects of personality and performance. He will assume too much skill on their part and fail to provide sufficiently explicit instructions.<sup>16</sup> It is also possible to fall into the opposite error and describe qualities and methods at such detailed length that the ordinarily intelligent rater is annoyed by the verbosity of the instructions and fails to study them carefully.

#### Pseudo-Sciences.

Pseudo-sciences are considered false doctrines. Opinions may be based upon pseudo-sciences that have little or no basis in fact and have been scientifically discredited. Judging a person's character by the shape of his

---

15. Laird, Donald A., op. cit., p. 104-8; Viteles, Morris S., Industrial Psychology, p. 171-8.

16. Kingsbury, Forrest A., op. cit., p. 379.



head or by his facial or physical features can accurately be called a method of fortune telling. Blonds do not conform to the so-called normal blond traits such as dynamic, driving, domineering, impatient, changeable, variety-loving, etc. Red haired people do not always have bad tempers. Handwriting is a very poor indicator of character traits. Neither do other physical characteristics indicate special character traits. Viteles says, "When submitted to examination under experimental conditions none of the character analysis systems has withstood the test".<sup>17</sup>

---

17. Viteles, Morris S., op. cit., p. 191-9. See also Laird, Donald A., op. cit., p. 109-38; and Dashiell, John Frederick, op. cit., p. 144-5.

## CHAPTER V

### RATING SCALE CRITERIA

## CHAPTER V

### RATING SCALE CRITERIA

Rating scales are measuring instruments. They may be useful, useless, or even detrimental. What distinguishes the useful rating scale? How is it identified? Some commonly accepted "earmarks" of a good rating scale are simplicity, brevity, convenience, clearness of directions, time required for completing the rating, agreeableness of the rating task, universality of the scale, and ease of scoring.<sup>1</sup>

However, the above points fail to insure the efficiency of a rating scale. It is commonplace to have rating scales appear in popular magazines and Sunday newspapers which claim to "rate your husband" or "wife", etc., which possess all of the above "earmarks" of good rating scales but which still fail as an accurate measuring instrument.

Rating scales must also objectively, consistently, and dependably measure what they are supposed to measure.<sup>2</sup> Scientific methods of checking these points have been developed. These will be discussed under the headings: objectivity, reliability, and validity.

Objectivity.

Objectivity in rating scales is defined as freedom

- 
1. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 135.
  2. Viteles, Morris S., op. cit., p. 201.

from personal opinion. To be objective, different users of equal ability must be able to apply a reliable rating scale to the same men and get the same results providing the raters are equally well acquainted with the ratee.<sup>3</sup> Also to be perfectly objective the same rater or different raters must score the scale in an identical manner today or a month hence.<sup>4</sup>

For instance a pair of scales are highly objective because different observers agree closely as to the weight of an article at any given time. Methods of statistically testing objectivity in rating scales follow:

1. Compare the scores of different raters on the same person. The different raters should agree closely providing they know the ratee equally well. However, the men to be rated may not exhibit the same behavior while in contact with the different raters due to different situations and differences in the raters. Therefore it is difficult to find two raters who have sampled identical ratee behavior and know the ratee well enough to rate him. When the raters are not judging the same thing it is no longer a test of rating scale objectivity but evidence of reliability. For this reason this method of testing rating

---

3. See Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 133; Laird, Donald A., op. cit., p. 83.

4. Bossing, Nelson L., Progressive Methods of Teaching in Secondary Schools, p. 641.

scales is usually classified as a test of reliability.<sup>5</sup>

2. Compare the total scores given at different times on the same set of ratings by one scorer. Or compare the results by different scorers on the same set of ratings. Practically perfect agreement should be expected in each case as a stencil is usually used in scoring.

### Reliability.

Reliability of rating scales is defined as the consistency with which the scale gauges the ability of those rated.<sup>6</sup> Another definition is the degree of accuracy with which the scale measures whatever it measures.<sup>7</sup> For example, a pair of scales are reliable because under the same weight they consistently register the same amount.

Kornhauser, of the University of Chicago, cautions blanket acceptance of low reliability coefficients as evidence that the scale is unreliable. He points out, "Traits like industry, accuracy, cooperativeness, and initiative are not necessarily permanent traits which appear under all circumstances and in all individuals' activities". He adds that disagreements between ratings of the ratee may indicate that the ratee is different under dif-

---

5. See Burt, Harold Ernest, op. cit., p. 245-253; and Paterson, Donald G., "The Scott Company Graphic Rating Scale", Journal of Personnel Research, Vol. 1, p. 361-376.

6. Yoder, Dale, op. cit., p. 269.

7. Bossing, Nelson L., op. cit., p. 644.

ferent circumstances--not that raters are wrong in their estimates.<sup>8</sup>

Methods of testing for reliability follow,<sup>9</sup> with the two most common listed first.

1. Re-rate the individuals after a short interval, an interval sufficiently brief so that the ratee cannot have changed greatly. Each rater's paired ratings should then be compared statistically to discover whether or not they are consistent. (Yoder says<sup>10</sup> that a coefficient of correlation of .75 is needed.)

2. Compare a ratee's standings on the first half of the scale with his standing on the second half of the scale or his standing on the odd numbered questions with his standing on the even questions. This is often called the split-half method.

3. Compare ratings on the same men by different judges.<sup>11</sup> If there is close agreement between ratings on the same men by different raters, who are equally proficient, it is evidence that the rating scale is reliable. However, it is recommended that one of the two methods

---

8. Kornhauser, Arthur W., "Reliability of Average Ratings", *Journal of Personnel Research*, Vol. 5, p. 309-17.

9. See Yoder, Dale, op. cit., p. 269-70; Laird, Donald A., op. cit., p. 3; Burt, Harold E., op. cit., p. 345-53; Walters, J. E., *Applied Personnel Administration*, p. 175; and Garrett, Henry E., op. cit., p. 311-24.

10. Yoder, Dale, op. cit., p. 270.

11. See Burt, Harold, op. cit., p. 348; and Paterson, Donald G., op. cit., p. 367.

already mentioned be tried first.<sup>12</sup>

4. Examine a distribution curve of the rating scores assigned by a rater and see whether the distribution is normal. This method assumes that the distribution of accurate measures of any ability in a sampling of the general population takes the form of the normal probability curve. However, the results are only an approximation as the men rated may not be a representative sample of the general population.<sup>13</sup>

5. Compare the theoretical mean and median with the actual average mean and median. This also assumes a normal distribution and is subject to the same inaccuracies that were pointed out in the last paragraph. Small groups are not expected to produce normal averages or normal distributions.<sup>14</sup>

6. Examine the distribution spread to see if sufficient discrimination is allowed to distinguish between one man's ability and another's.<sup>15</sup>

#### Validity.

Validity of rating scales is defined as the fidelity with which the scale measures what it is intended to

---

12. Clinton, R. J., in a conversation with the writer, April 18, 1939.

13. See Burt, Harold E., op. cit., p. 345; and Bingham, Walter V., and Freyd, Max, op. cit., p. 133-4.

14. Probst, J. B., op. cit., p. 44.

15. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 134.

measure.<sup>16</sup> If the ratings are valid, the man who is rated A and the man who is rated E must be in fact A and E men.<sup>17</sup>

A measuring device may be reliable, that is, agree closely with itself, but still be lacking in validity. A pair of scales may indicate consistently that a man weighs 185 pounds but if he really only weighs 160 pounds, the weight indicated is wrong and the scales are invalid.

Validity is the very essence of the rating plan, yet it is the most difficult factor to prove. The most common method of checking the rating scale is to compare the ratings with some criterion of success. However, Burt<sup>18</sup> says that it is often impossible to obtain an accurate production criterion of success with which to compare the rating scale results.<sup>18</sup> Methods of checking validity follow.

1. Compare or correlate the ratings with some objective measure of success such as freedom from errors for a typist or cords of wood cut for a wood cutter.

2. Compare the group's rating with other criteria consisting of a consensus of opinion, based upon the best available independent estimates, of the success of the employees. The independent judgments are often in the form of classifications into rank orders.<sup>19</sup>

3. Compare ratings with test scores. Ratings on

---

16. Bossing, Nelson A., op. cit., p. 641.

17. Probst, J. B., op. cit., p. 52.

18. Burt, Harold E., op. cit., p. 353.

19. Probst, J. B., op. cit., p. 52.



intelligence might be compared with intelligence test scores.<sup>20</sup>

4. Correlate the standings of unrelated abilities as shown by a rating. If the correlation between appearance and intelligence is high it may indicate low validity. This method is used in determining the influence of halo.<sup>21</sup>

5. Have a friend of the ratee identify him from a set of ratings. The scale which yields the most easily identified ratings is best.<sup>22</sup>

6. Obtain the judgments of competent persons as to the validity of the rating scale.<sup>23</sup>

7. Compare the qualities or traits covered in the rating scale with the accepted qualities needed for success in that field. If the scale includes the qualities in their true relationship as agreed by authorities it may be said to be high in validity.

The task of evaluating rating scales is exceedingly complex. There are intangible and unmeasured effects of ratings which are not appraised statistically. These include such things as the effects on morale or the growth of personnel spirit among raters. Kornhauser states, "Technical statistical studies alone can never give a final

---

20. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 132.

21. Ibid., p. 134.

22. Ibid., p. 135.

23.

judgment as to the value of the effects produced by rating scales".<sup>24</sup>

---

24. Kornhauser, Arthur W., "What Are Rating Scales Good For?", Journal of Personnel Research, Vol. 5, p. 193.

## CHAPTER VI

### REACTIONS TO RATING SCALES

## CHAPTER VI

### REACTIONS TO RATING SCALES

Although the common purpose of all rating scales is to obtain more systematic and accurate evaluations,<sup>1</sup> so much difficulty has been experienced in securing accurate ratings that many personnel managers despair of their usefulness.<sup>2</sup> They have preferred to trust the judgment of superior officials. "On the other hand, it is pointed out that these officials must use some kind of a standard in evaluating the usefulness of the subordinates, and that every effort should be made to eliminate irrelevant items from their judgment and to objectify so far as possible their conclusions", according to Leonard D. White, Professor of Political Science at the University of Chicago.<sup>3</sup>

Dale Yoder, Professor of Economics and Industrial Relations at the University of Minnesota, says, "It is probable that no personnel practice is more constantly utilized and at the same time more generally unsatisfactory than service rating. Management everywhere, both in large concerns and in small, looks to ratings as an important tool, and almost everywhere, when ratings are carefully analyzed, they are found to be highly questionable in validity."

---

1. Watkins, Gordon S., op. cit., p. 325.

2. White, Leonard D., op. cit., p. 312.

3. Ibid.

4. Yoder, Dale, op. cit., p. 246.

White states that the feeling remains widespread, "... that service ratings are not a fair reflection of the real usefulness of the employee which is alleged to be far too subtle a thing to be caught within the four corners of any formal rating".<sup>5</sup>

In 1921 C. W. Reed, at that time a member of the California Commission, announced that he had studied efficiency records in Chicago, New York, and Washington, and had found them nowhere a success. He also added that he saw no future for them.<sup>6</sup>

"Mr. William C. Beyer, Director of the Bureau of Municipal Research of Philadelphia, wrote of service rating systems that 'none so far has stood the test of time; all except the most recent experiments, which have not been in operation long enough to be judged, have proved disappointments'".<sup>7</sup>

Mr. Lucius Wilmerding, Jr., Assistant to the Director of Research at the Commission of Inquiry on Public Service Personnel, concluded, "Notwithstanding the seeming scientific accuracy of efficiency-rating systems, none of them has proved successful in practice".<sup>8</sup>

---

5. White, Leonard D., op. cit., p. 317.

6. Transactions, Commonwealth Club of California (1921-22), p. 323.

7. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 428.

8. Wilmerding, Lucius, Jr., Government by Merit. Commission of Inquiry of Public Service Personnel (New York, 1935), p. 169.

Despite these pessimistic observations, as the former Bureau of Public Personnel Administration pointed out, the problem is one "which simply will not down either in the public service or in large commercial establishments".<sup>9</sup>

Mosher and Kingsley point out, "The reason for this is that in one way or another employees are being rated and must be rated from time to time in the ordinary conduct of business. Rating is, therefore, a practical and basic problem". They add that its solution presents a problem of the first magnitude.<sup>10</sup>

In answer to the question, "If rating systems do not have a mathematical precision and are largely records of personal opinion, why use them?", Watkins replied, "Because in the final analysis modern industry and business rest upon subjective personal judgments. The successful executives are those whose personal opinions produce valuable results. But best results are obtained when personal judgments are supplemental by the more objective evidences such as records of attendance, performance, and cooperativeness. Despite their obvious shortcomings, rating systems have marked advantages."<sup>11</sup>

Bingham, Director, Personnel Research Federation, and

- 
9. What's Wrong with Service (Efficiency) Ratings? Public Personnel Studies, Vol. 7, 1929, p. 18.  
10. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 429.  
11. Watkins, Gordon S., op. cit., p. 327.

Freyd, Research Associate, Personnel Research Federation state, "The records (service records) represent subjective impressions and do not have the objectivity and reliability of test scores and other more accurate measures of abilities. Nevertheless, they are an improvement over unsystematic judgments."<sup>12</sup>

The opinion of Tead and Metcalf is, "The fact that the technique of rating is still highly experimental does not alter the fact that any conscientious and explicit effort to state and measure relatively the desirable and undesirable qualities in a supervisor can almost certainly bring to light deficiencies, errors in judgment, prejudices, limitations of capacity, that formerly were unsuspected and in consequence left undisturbed".<sup>13</sup>

Although rating methods have been subjected to criticism and many of them thrown into the discard, it can hardly be denied that significant progress has been made and, surely, if the alternative is to fall back on the untrammelled opinions of "human, all-too-human" supervisors and executives, it is better to utilize even imperfect instrument and seek to improve it.

White holds out a strong incentive when he says, "The

---

12. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 122.

13. Tead, Ordway, and Metcalf, Henry C., Personnel Administration--Its Principles and Practice, p. 165.

14. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 432.

invaluable aid which a sound service rating could render, directly or indirectly, to nearly all phases of personnel management impels constant efforts to construct better systems than have yet been made available".<sup>15</sup>

Mosher and Kingsley summarize the situation with, "It is frankly recognized that rating systems are rather crude and imperfect means of estimating and recording abilities and habits of work. But since they are better than scattering and incomparable judgments of individual executives, the personnel division should accept the challenge of the situation and seek to provide more and more adequate and useful instruments."<sup>16</sup>

---

15. White, Leonard D., op. cit., p. 319.

16. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 443.



## **CHAPTER VII**

### **EARLY RATING METHODS AND DEVELOPMENT**

## CHAPTER VII

## EARLY RATING METHODS AND DEVELOPMENT

Early in the nineteenth century Robert Owen installed a rating system in the New Lanark Cotton Mills in Scotland. This was one of the first rating devices and it consisted of "character-books" and character "blocks". Each operator had a square block of wood placed by the side of his machine or bench. The four sides of this block were painted black, blue, yellow, and white respectively. The black stood for bad conduct; blue, indifferent; yellow, good; and white, excellent. Each morning the foreman would arrange the block according to the conduct of the employee on the previous day. The results were recorded in the character-book, which was always consulted upon promotion. It was possible for Mr. Owen to tell at sight the conduct of each operator by these "silent mentors" as he called them as he passed through the shops.<sup>1</sup>

Early literature of scientific management recognized but did not develop the idea of rating employees. Taylor, in 1895, wrote with reference to the Midvale Steele Works, "A careful record has been kept of each man's good points as well as his shortcomings, and one of the principal duties of each foreman was to make this careful study of his men so that substantial justice can be done to each".<sup>2</sup>

---

1. Yoder, Dale, op. cit., p. 250; and Walters, J. E., op. cit., p. 42-43.

2. White, Leonard D., op. cit., p. 312.

The principle that promotions should be made only upon satisfactory evidence of efficiency was proposed in 1906 by F. H. Hitchcock, First Assistant Postmaster General. He said, "To avoid any injustice in reporting efficiency the records should be as definite in nature as conditions will permit". The system was installed in all city delivery post offices by 1909, and provided for ratings of clerks and carriers as to speed and accuracy, conduct, attendance, and knowledge. Knowledge was measured by examination while the "best man" was used as the standard for the other items.<sup>3</sup>

The World War served as a great impetus to the use of rating scales and Walter Dill Scott developed the Man-to-Man comparison scale for use in the United States Army. Since that time many scales have been developed. A wide variety of schemes have been propounded and experimented with. These range from a simple method where numerical percentages of letter (A to F) are assigned on a kind of over-all basis to the grading of dozens of traits and behaviors, all of which may be combined into a single score by an elaborate formula for weighting purposes. The first is too simple and invites so many inconsistencies as between individuals grading the same employee that it has been thoroughly discredited.<sup>4</sup> On the other hand, the

---

3. Ibid., p. 313.

4. Probst, J. B., (Foreword by Fred Telford), op. cit., p. 5-8.

latter may be so complex, obscure, and difficult to handle that it falls of its own weight. Within the range set by these two extremes may be mentioned the following: check-list rating scale, numerical rating scale, alphabetical rating scale, adjective rating scale, descriptive rating scale, multiple step rating scale, linear rating scale, order of merit scale, Scott Man-to-Man Rating Scale, graphic rating scale, Probst Service Rating Scale, and the Service Rating Form used in the Forest Service.

Considerable progress has been made since the days of the school-grading schemes applied on an over-all basis. This progress has been particularly marked by, first, the utilizing of specific desirable traits in different classes of workers, from laborers to executives; second, objectifying the terms used in describing these traits, as, for example, the trait of leadership: "capable and forceful leader", "handles work well", "fails to command confidence", etc.; third, by training rating officers; fourth, by efforts to adjust ratings according to a norm; and, fifth, by working out better scoring methods.<sup>5</sup>

---

5. Mosher, William E., and Kingsley, J. Donald. Public Personnel Administration. New York, Harper and Brothers, 1936, p. 432.

## **CHAPTER VIII**

### **TYPES OF RATING SCALES**

## CHAPTER VIII

### TYPES OF RATING SCALES

The generally known rating plans will be briefly described and discussed.

#### Check-List Rating Scale

According to Viteles this is about the simplest type of rating scale.<sup>6</sup> In it the rater is required to give his judgment in the form of a "yes" or "no" answer to the question: "Does the man possess this ability?" Thus a check list of abilities is supplied and the rater makes the appropriate answer. The check list may ask:

1. Is the rates dependable?
  2. Is he industrious?
  3. Is he honest?
  4. Is he quarrelsome?
  5. Does he work independently?
- Etc.

This method gives some definite terms with which to work and the rater doesn't have to grope for words with which to express himself. However, this method raises two serious difficulties. It assumes that a person either has "all-or-none" of any quality, that he is either wholly un-dependable or wholly dependable. It is obvious that human traits are present in different persons in varying degrees.

---

6. Viteles, Morris S., op. cit., p. 211.

not either absent or present. Abilities are usually distributed among people in general in accordance with the normal probability curve.<sup>7</sup> Professor Thorndike has stated, "Whatever exists, exists in some amount".<sup>8</sup>

Also, it fails to discriminate between workers. Both John and George may be judged dependable, honest, and industrious, in which case the final ratings will be the same.

Bingham and Freyd entitle this method, "The scale of alternatives or two steps".<sup>9</sup> It also goes under the name "yes and no" type rating scale.

#### Numerical Rating Scale.

In this scale the rater judges each man in terms of a number of percentage in much the same way as school examination papers were ordinarily graded. This scale takes the following form:

Qualification	Rating (1 to 100 per cent)
Performance	_____
Leadership	_____
Character	_____

A rating of 100 per cent or 10, for example, assumes per-

7. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 126.

8. Laird Donald A., op. cit., p. 180-81.

9. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 126.

fection on the trait being rated, whereas, 10 per cent or 1 implies that the man does not possess the particular trait under consideration.

Difficulty is encountered in getting raters to be consistent. Two raters can rarely be found who agree on how much 100 per cent of the ability is, or 80 per cent, or 50 per cent.<sup>10</sup> What is satisfactory and what is unsatisfactory? If it is set at any other point than 70 per cent it is difficult to get away from the habit built up by long years of association in school that 70 per cent is the dividing line between satisfactory and unsatisfactory work. Moreover, this method implies a finer discrimination than is possible or desirable. Such scores appear to be highly accurate since one person may get 86 and another 87. However, "Experiment has indicated that it is almost impossible for anyone to estimate these traits with accuracy greater than is allowed by a scale of from 1 to 5".<sup>11</sup> Therefore this mathematical precision gives a false sense of security.

#### Alphabetical Rating Scale.

This is essentially the same method as Number 2 (numerical), but letters, instead of numbers, are employed to distinguish degrees of excellence on the trait in

---

10. Ibid.

11. Laird, Donald A., op. cit., p. 182.



question. For example, A or E may stand for excellent; B or G, good; C or F, fair; and D or P, poor.<sup>12</sup>

The too fine discrimination objection raised above is removed but the other objections remain including the stigma of school grades.

Adjective Rating Scale.

An example of this type follows:

Abilities	Poor	Fair	Good	Very Good	Excellent
Concentration					
Observation					
Imagination					
Memory					
Judgment					
Initiative					
Will Power					
Self Confidence					

Instead of letters or numbers, such words as excellent, very good, good, fair, and poor are used to indicate the

---

12. Viteles, Morris S., op. cit., p. 211.

rating. Various numbers of classes are used, usually ranging from three to nine.<sup>13</sup>

The ratings do not give such a fictitious accuracy to the system as the numerical scale does. However, it still has its weaknesses. There is no way to be sure that excellent in initiative means the same thing to two different raters. Also they may be rating different qualities even though it is entitled initiative. In each case the quality of initiative may have different meanings to different raters. To one it may mean industry and to another, self-reliance. In other words they may disagree as to the meaning of the term.<sup>14</sup>

#### Descriptive Rating Scale.

Descriptive adjectives and phrases are used to describe different grades of ability from one extreme to another in this type of scale. This is done for each trait that is considered essential by the management. "Here is an example of such a list for rating one aspect of social behavior:

1. Extremely breezy and informal
2. Cordial and congenial
3. Meets one half-way
4. Slightly reserved

---

13. Baridon, Felix E., and Loomis, Earl H., op. cit., p. 190.

14. Laird, Donald A., op. cit., p. 181-3.

### 5. Constrained and formal."<sup>15</sup>

The rater assigns to each man the number corresponding to the phrase which describes him best.

This type of rating scale is an attempt to make the steps more definite and concrete than in the usual characterizations of a man's ability from excellent to poor.

#### Multiple Step Rating Scale.

This method is very similar to the adjective rating scale, already described. In this method symbols are used instead of words. In the following examples Y represents yes and N represents no in answer to the question: "Does the man possess this ability?".

Does the employee possess the ability  
to conduct himself tactfully? Encircle  
the symbol which represents your answer.

Y!      Y      ?      N      N!

An alternative form is:

Y      y      ?      n      N<sup>16</sup>

Plus and minus signs are also adaptable to this method.

This avoids undesirable associations which may be present when words such as good, fair, excellent, etc. are used although it retains other weaknesses, particularly, the possibility of rater disagreement as to the precise amount of any quality denoted by any one of the terms in

---

15. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 127.

16. Ibid.

the rating scale as well as disagreement as to the meaning of the abilities to be rated.

### Linear Rating Scale.

In this scale a straight line is drawn to the right of each trait to represent the range of ability, one end representing the least amount and the other end the greatest amount. A man is rated by making a check along this line at the place which corresponds to his standing in that ability.<sup>17</sup> A sample of this type of rating scale follows.

Physical condition

Sickly \_\_\_\_\_ Robust

Another variation follows:

Personal Qualities

Poor \_\_\_\_\_ Excellent

Common sense, industry, dependability, judgment, initiative, force, self-reliance, punctuality, courtesy, temperament, sense of humor, freedom from conceit and selfishness, readiness and ability to cooperate, etc.

This rating scale overtaxes the powers of discrimination of raters and it is very difficult to get intelligent agreement or disagreement in ratings due to its subjectivity. Its susceptibility to averaging of ratings and to the halo effect will be discussed under the graphic rating scale where the weakness occurs in the same manner.

---

17. Ibid.

Order of Merit Rating Scale.

This is the first of two forms in which the reporting is done with men rather than with words. In this scale the rater, considering each ability in turn, arranged the group of men in order of merit, heading his list with the member of the group having the greatest amount of the ability in question and ending with the member having the least amount. After the men in the group have been arranged in rank order they may be given a number for their standing. For example, assume that 20 men were rated. The highest might receive a score of 20, the second a score of 19, etc. This method of scoring has the disadvantage of making records from different departments incomparable unless there are an equal number of employees in each department.<sup>18</sup>

A scoring procedure that eliminates this disadvantage is to divide the ranked employees, regardless of how many there may be into fifths. If there are 20 employees being arranged in this way, the highest four would be given a score of 5, indicating that they were in the highest fifth. The next four would be given a score of 4, indicating that they were in the second highest fifth. Similarly in case the number of employees reported on was sixty, the highest 12 would receive a score of 5, the next twelve a score of 4, and so on. The numerical values are determined by the

---

18. Laird, Donald A., op. cit., p. 184-5.

group itself and the reports of one executive can be made comparable with those of another.<sup>19</sup>

Arranging men in a rank order is a brain-racking process, but there is a simple way of making the task less disagreeable. Each man's name is written on a card. The rater sorts these cards into three groups, representing men who are high, average, and low in the ability in question. Then he puts the men in each group in rank order. By assembling the groups in proper order the ranking is completed.<sup>20</sup>

The order of merit method has the advantage of getting at the crux of the rating problem by disregarding all other persons or impressions and requiring comparisons between the members of the group under consideration.

However, it also has some distinct disadvantages. It makes distinctions in ability which are impossible to make in a large group. It gives the rater no framework on which to hang his judgments and it assumes that a unit difference in rank always indicates a constant difference in abilities, unless statistical corrections are made. As a matter of fact this difference is greater at the extremes of ability than in the middle range. In other words, if 60 men are ranked in any ability, the difference in

---

19. Ibid.

20. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 123.

ability between the first and second man is greater than the difference between the twenty-fourth and twenty-fifth man.<sup>21</sup>

Scott Man-to-Man Comparison Scale.

Walter Dill Scott originally developed this rating method at Carnegie Institute of Technology for the selection of salesmen.<sup>22</sup> As Director of the Committee on Classification of Personnel in the United States Army he modified that scale for use in rating the efficiency of army officers and it became the first rating scale to be used to any great extent. This scale supplanted the seniority system of promotion in the army and initiated an era of promotion on the basis of merit.<sup>23</sup>

This scale contains five subjects for consideration--physical qualities, intelligence, leadership, personal qualities, and general value to the service. The rater makes a master scale for each ability in the group (or among his acquaintance), a man of average ability, and two other men, one midway between the highest and the average man and one midway between the lowest and the average man. These five men serve as standards by which to judge all the other members of the group with reference

---

21. Ibid., p. 124.

22. Ibid.

23. Paterson, Donald G., "The Scott Company Graphic Rating Scale", Journal of Personnel Research, Vol. 1, p. 361.

to the ability in question. Arbitrary numerical values are set for each grade. For example the highest man may receive 25 points, the high man 20 points, the middle man 15 points, the low man 10 points, and the lowest man 5 points. A sample master scale for physical qualities looks like this:

PHYSICAL QUALITIES	Highest	<u>William Row</u>	25
Physique, bearing, neat-	High	<u>James Smith</u>	20
ness, voice, energy, en-			
durance.	Middle	<u>John Doe</u>	15
Consider how he impres-	Low	<u>John Edwards</u>	10
ses his command in these			
respects.	Lowest	<u>George Jones</u>	5

The rater prepares a master scales in this way for each of the five subjects and the rater estimates the ability of the person being rated in comparison to those on his master scale of individuals. If he thinks the person is about the same in that ability as Row, the highest in the scale, he gives him 25 points. If he compares to the lowest man, Jones, he receives 5 points, and so on. Finally a total for all characteristics is made.<sup>24</sup>

This is a somewhat cumbersome procedure although this scheme marked a real advance over earlier experiments. It objectifies judgments by rating each man in comparison with known men instead of rating each man on intangible factors.

Usually the rater will not give the proper time and

---

24. Burt, Harold Ernest, op. cit., p. 329-334.



thought to the making of the master scale. Besides difficulties in selecting men for the master scale, there are often variations in these selections by different supervisors and others rating the same group.<sup>25</sup> These variations in yardsticks by different raters may result in one rater's highest being another's middle, and his lowest may be another's low. Some men tend to ignore the key men entirely in making their ratings and fall back upon their general conceptions of excellence in the ability in question.<sup>26</sup> These difficulties have resulted in a gradual disuse of this type of rating scale.

Graphic Rating Scale. (Also called Sky Line or Profile Rating Scales)

This is one of the most used rating scales. It was originated in 1920 by members of the Scott Company, Philadelphia to remedy defects in the army scale.<sup>27</sup> Credit for its origin as well as supervising its experimental development belongs to Beardsley Ruml.<sup>28</sup>

The graphic rating scale is a combination of the straight line rating scale, representing the range of the ability, and descriptive rating scale, placed beneath the

---

25. Viteles, Morris S., op. cit., p. 210.

26. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 126.

27. White, Leonard D., op. cit., p. 314.

28. Paterson, Donald G., op. cit., p. 131.

line to denote certain points in the range.<sup>29</sup> This is done for each ability and results in an extension of a form similar to the following.

APPEARANCE.

Consider neatness of person and dress.	Appro- priate	Neat	Ordin- ary	Pass- able	Slov- only
-------------------------------------------	------------------	------	---------------	---------------	---------------

ACCURACY.

Consider quality of work, freedom from errors.	No errors	Very careful	Few errors	Care- less	Many errors
------------------------------------------------------	--------------	-----------------	---------------	---------------	----------------

The rater is instructed to indicate with a check mark the point on the line corresponding most nearly to the degree of the qualification found in the employee who is being rated. The ratings can later be expressed as numerical values and they are usually scored with the aid of a calibrated scoring stencil. Occasionally lines are drawn connecting these checked points, hence the origin of the names graphic, profile, and sky-line.

Several advantages have been claimed for this method. The rater isn't limited to five steps, but can make as fine gradations in judgment as he cares to. At the same time the phrases are there as aids in making the judgments definite and concrete. Laird points out that it is easy to use, not cumbersome as is the man-to-man scale, and that it is much more definite because it is a scale of human behavior rather than vague adjectives such as aver-

---

29. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 131.

age, high, and low.<sup>30</sup>

Probst points out a number of weaknesses in the graphic scale. Some of Mr. Probst's criticisms are aimed at specific forms of graphic scales and cannot be sustained for all types. Among those that he mentions are;<sup>31</sup>

1. Too few factors. Many of the graphic scales have only three or four factors such as quality of performance, productiveness, and qualifications which each include several traits or qualities. It is too difficult to consider several traits in detail at once and give a single accurate grade. This criticism is largely eliminated in many forms by judging single traits and thereby lengthening the scale.

2. Manner of reporting. It is futile to try and get reporting officers to do the refined marking expected.

3. Tendency to "average" ratings. It is contended that because the rater is forced to judge all points, even though it is constructed in such a manner as to require the rater's attention on a single aptitude at a time, that he will lack specific information on many points. In the absence of sufficient intimate knowledge of the employee he rates him "average". Also because of doubt as to where his check mark should go the rater tends to gravitate towards the center.

---

30. Laird, Donald A., op. cit., p. 189.

31. Probst, J. B., op. cit., p. 17-18.

4. Halo effect. "With the graphic rating plan, the reporting officer, when influenced favorably by 'halo', tends to place all his check marks near the left end of the line--the point of maximum excellence." It would be equally true that the negative influence of "halo" would tend to place all check marks near the opposite end, the point of minimum efficiency. Some attempt has been made to meet this criticism by reversing the order of excellence, that is, make the right end of the line the highest part of the time. Regarding this Probst says, "This theory would...be at best not only confusing but highly irritating and annoying to the reporting officer".<sup>32</sup>

5. Adjusting the ratings. Adjustments must be made when graphic rating scales are used in order to bring about a proper degree of consistency between high and low raters. Some reporting officers rate all employees higher than they should, while others rate lower than they should. An example is cited of rating secured by the graphic method in the city of St. Paul showing the absence of satisfactory results even with the adjustment of ratings. One fire station, generally known to be among the lowest in efficiency, had an average rating for the station of 94, while one of the best stations showed an average rating for the station of 85.<sup>33</sup>

---

32. Ibid., p. 19.

33. Ibid.

Probst Service Rating Scale.<sup>34</sup>

This recent devise for determining service value was developed by Mr. J. B. Probst, the chief examiner of the St. Paul Commission. The scale contains a list of about one hundred modes of behavior or characteristics, the majority of which are objectively observable. The reporting officer is expected to check the true facts concerning each individual employee and is not to consider the employees as a group even for comparison. Most of the characteristics are qualities that are either above or below the average and are outstanding. This is done so that rating officers may be better able to check some of the items with assurance. Raters check only those items which are known to describe the ratee and disregard any traits or characteristics of which they have incomplete knowledge. Preceding the columns of characteristics are three columns of blocks providing for a report by three superiors of the employee, if that number is acquainted with his work. It is expected that the rater lowest in authority shall check the employee first, someone in higher authority next, and the rater highest in authority last. Rating officers are encouraged to compare their marks with those of the officers who have already checked.<sup>35</sup> A copy of this rating scale<sup>36</sup>

---

34. Mosher, William E., and Kingsley, J. Donald, op. cit., 435-440.

35. Probst, J. B., op. cit., p. 25.

36. See bottom page 62.

may be found in the appendix.<sup>37</sup>

Although the reporting officers report or check the items on known facts concerning the employee, the actual rating of the employee is obtained through a separate evaluation of those facts, generally conducted by the central personnel agency, by means of a stencil and computing scale.<sup>38</sup> The scoring and rating system was developed through numerous experiments and studies of the distribution of ratings under various scoring procedures and through the employment of the normal curve and other criteria, similar to those employed in the evaluation of tests.<sup>39</sup>

The Probst system has succeeded in "translating more or less general and complex traits into behavioristic or quasi-behavioristic terms. Behavior can be observed and agreement among several raters is likely to result. The

---

36. There are now seven report forms available from Probst Rating System. They include general for several classes of service including clerical, general maintenance, general inspection, salesmen, etc.; professional for such classes of service as engineering, medical, library, social service, etc.; educational for instructors and teachers; police; fire, for firemen; general or labor for skilled and semi-skilled labor; labor for unskilled labor; and personality for applicants.

37. For an elaboration of this approach see Probst, J. B., Service Ratings, Baltimore, Lord Baltimore Press, 1931.

38. Brumbaugh, H. C., Report on Probst Rating System by Multnomah County Civil Service Commission, (mimeographed, Dec.) 1935, p. 6.

39. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 440.

formulation of many such terms in striking and arresting phrases is also a contribution of real value."<sup>40</sup> Doing the actual rating and scoring in a central agency will make for much-to-be-desired uniformity, if suitable scoring methods can be devised. This presupposed that they can be checked and reviewed by any interested party.<sup>41</sup>

Mr. James McCarran, secretary of the Multnomah County Civil Service Commission, says that he considers the Probst rating system the best rating scale out, the most objective. He added that the chief trouble with this scale, which is true of any scale, is to get honest rating and that the grades are too high. The raters need some education regarding making service ratings to the actual rating.<sup>42</sup>

Mr. H. C. Brumbaugh, chairman of the Multnomah County Civil Service Commission states:<sup>43</sup>

"The Probst system provides the following, which cannot be obtained through any other known rating system:

1. That the employee's performance is reported not merely in general conclusions as an expression of opinion, but in statements of fact or specific and verifiable judgment.
2. The facts, traits or qualities are stated on the report form in terms of the everyday thinking of the reporting officers and not in letters or percentages.
3. The reported facts can be scientifically interpreted and evaluated by a proved process

---

40. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 440.

41. Ibid.

42. McCarran, James. Stated during a conversation with the writer, March 17, 1939.

43. Brumbaugh, H. C., op. cit., p. 6-7.

based on thoroughly sound principles.

4. The report sheet permits the reporting officer optional selection of various traits and qualities so that he may report on only those things of which he has definite knowledge.

5. The reporting officer is not required to measure relative degrees of quality in different employees.

6. The rating and scoring system used makes it unnecessary to adjust the resulting scores or ratings to compensate for high or low raters.

7. The reporting officer is virtually forced to report accurately or be shown by the evidence in his own reports not to have done so.

8. The scoring system is sufficiently simple for the employees to understand and determine for themselves in a general way the fairness and reliability of the rating."

Mr. Probst, the originator of the system, adds another point which he decided the new system must provide, as far as possible:

"That the scheme also take into consideration the ordinary mental processes of the reporting officer--his reluctance generally to rate negative qualities, his normal desire to say good things about an employee, his tendency to use superlatives in describing the favorite employee, and the like."<sup>44</sup>

Some are not ready to acknowledge that the system in actual operation achieves all of these ends,<sup>45</sup> but Mosher and Kingsley say that "in general,.....the plan has been more satisfactory than earlier schemes and is worthy of extended trial".<sup>46</sup>

---

44. Probst, J. B., op. cit., p. 23.

45. Blog, Leon, "Does the Probst Rating System Rate?", National Municipal Review (1931), Vol. 20, p. 581.

46. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 436.



However, this system is far from perfect and further experimentation is distinctly called for. In the first place, too much is claimed for the scale.<sup>47</sup> Specifically the two most serious criticisms seem to be:

1. The advisability of three raters checking each employee on the same sheet is questioned. Mr. Probst argues that independence is achieved by having the rater highest in authority rate last. This assumed that the rater in highest authority is equally as conversant with the abilities and disabilities of the ratee as the immediate supervising officer and this is often untrue. It is quite possible that the later raters, realizing that the immediate supervising officer is more familiar with the ratee, will tend to follow his lead. Also the fact is overlooked that the immediate supervising officer may be influenced in his checking by the known likes and dislikes of his boss, if that individual is to go over and re-check every service report.<sup>48</sup>

2. The Probst scale makes no provision for special qualities required in different classes of employment and certain qualities that are included are not called for in certain position. For example, the general form shown in the appendix states that it is "for appraising the service value of employees, supervisors or officers, other than

---

47. Ibid., p. 440.

48. Ibid., p. 439.

those in the police, fire, labor or educational services". Some qualities may be included for appraisal in this form which are not called for in a special type of work such as the profession of forestry. Neither is there any assurance that this scale includes and measures the special qualities and abilities needed in a forester according to their relative importance. This objection might be carried further because foresters are engaged in many types of work with special qualities needed in the various types of service such as research, district rangers, supervisor's staff, etc. In other words, one questions whether it hits the essential points and properly weighs them in all positions.<sup>49</sup>

Service Rating Forms Used in the United States Forest Service.

Service Rating Forms 3200 and 3201 were put into use in the Washington Office of the Forest Service in 1935 and extended to the field service in 1938. Prior to the inauguration of these forms, no standardized system of rating employees was in effect in the field service, other than for clerks. The various regions developed their own methods and each attempted to rate its employees as uniformly and fairly as possible.<sup>50</sup> William C. Hull, Executive

---

49. Ibid., p. 440.

50. Keplinger, Peter, Chief, Division of Personnel Management, U. S. Forest Service, in a letter to the writer, dated April 17, 1939.

Assistant of the U. S. Civil Service Commission states, "The original authority for rating employees on the basis of efficiency is contained in Section 9 of the Classification Act of 1923--however, this act applied only to civilian positions within the District of Columbia; therefore, the U. S. Civil Service Commission's jurisdiction in the matter of efficiency ratings is also limited to the departmental services, leaving within the discretion of the heads of the departments and establishments its application to the field service".<sup>51</sup>

The Service Rating Forms are a refinement of the numerical rating scale. Ratings of from 1 to 10 are given on each of three points: quality of performance, productiveness, and qualifications shown on the job. Each of these three points is further described and defined by from four to six general phrases indicating the traits included. Samples of traits mentioned are thoroughness, accuracy, neatness, industry, speed, initiative, resourcefulness, etc. A copy of each of the two forms used, Forms 3200 and 3201, may be found in the appendix.

In order to assist the rating officer in making his rating he is told to underline the elements which are especially important in the position and to mark the employee

---

51. Hull, William C., Executive Assistant of the United States Civil Service Commission, in a letter to the writer, dated Feb. 13, 1939.

on each element in one of three ways: if neither strong nor weak point, if weak point, or, if strong point.

The final score is found by adding the three ratings. This is converted into a final five class rating using the words excellent, very good, good, fair, or unsatisfactory. The sum of the ratings is also converted into another numerical score which has 93 as the highest possible score. This is the score that is used in the personnel records.

Raters using this rating scale report that because of its brevity it is quickly used, usually consuming one minute or less in marking. It is also easily and quickly graded. Performance on the job, which is what is important, is emphasized.

One rater says that this scale "gives an over-all picture of the whole"<sup>52</sup> person being rated. In other words the rating tends to express a summary<sup>53</sup> of the total pattern of the individual. This is an important point and Trow states it by saying, "It is the total pattern of the personality which counts, how the various characteristics harmonize with each other which determines the effectiveness of the (man being rated), not their arithmetic sum, and one discordant element might upset the whole balance. A squeaky voice and a funny face might ruin the success of a serious-

---

52. Kearns, R. S., in comments made regarding the Service Rating Form on January 6, 1939.

53. Starker, T. J., in comments made regarding the Service Rating Form on January 13, 1939.

mined domineering personality, but might be quickly forgotten or a real advantage to a cheerful, humorous, friendly individual."<sup>54</sup> Therefore it is particularly important that the whole person be judged upon his success in his work and the Service Rating Form used seems to offer that opportunity.

However, the rating scale used in the Forest Service has not been entirely satisfactory. "Numerous studies have been made on rating systems--in Washington but as yet nothing worth while has been developed", according to Mr. Paine. However, many recommendations for improvement have been made and indications are that it will be revised as soon as time permits.<sup>55</sup> Perhaps the most serious criticism that might be made is that adequate provision is not afforded to avoid the personal element, that the scale is not objective.<sup>56</sup> Disagreements are likely to occur in two ways. Raters are apt to disagree as to the exact meaning of the items to be rated as they are stated in more or less general, complex, and vague terms and most certainly not in a clear, concise, observable form of human behavior. In other words raters are likely to disagree as to the precise

---

54. Trow, William Clark, "How Shall Teaching Be Evaluated?", Educational Administration and Supervision, V. 20, p. 264-72.

55. Paine, P. L., in a letter to the writer, dated Feb. 22, 1939; and Keplinger, Peter, in a letter to the writer, dated April 17, 1939.

56. Mason, E. G., in comments made regarding the Service Rating Form, January 11, 1939.

meaning of such terms as "acceptability of work; thoroughness" as it affects "quality of performance" as long as they are stated in such general terms and in consequence the raters will be judging different things.

Second, there is apt to be rater disagreement in the value of the 1 to 10 ratings, or the corresponding words, excellent, very good, good, fair, and unsatisfactory. Does a score of 5 or 8 mean the same thing to two or more different raters? There is no assurance that it does.<sup>57</sup> The presence of this second problem of inconsistency is illustrated by an example given by a forest officer who told of two raters who turned in reports with the qualities checked identically with reference to the strong, average, and weak points on the left but who had different scores on the right. An attempt has been made to rectify this weakness by rating each of the elements on a graphic scale and the score for each set of elements is taken directly from the graphic scale which is based on the individual elements. This would insure identical scores if the elements were marked the same.

In addition to the lack of consistency between different raters in the use of the number or corresponding word ratings, such numbers or words are not consistently used by the same rater at different times.<sup>58</sup>

---

57. Laird, Donald A., op. cit., p. 181-3.

58. Yoder, Dale, op. cit., p. 266.

As a result of this lack of objectivity and behavior analysis, only a general opinion of the ratee can be given. In practice this general opinion rating is probably based on whether or not the rater considers the ratee eligible for a salary increase or not. This is unavoidable as long as the rater determines the total score after rating a subordinate, and the significance of this total score in terms of possible salary changes remains on the rating form.

Other criticisms include:

1. Too few factors. The discrimination of an individual is overtaxed when he is asked to analyze, compare, and evaluate a number of qualities in detail at once and give a single judgment rating for the whole as is expected in each of the three parts of this rating form.<sup>59</sup> This condition results in rater criticism such as, "too general--hard to grade objectively",<sup>60</sup> "elements in very large groups",<sup>61</sup> "each item covers a lot of ground",<sup>62</sup> and "hard to understand".<sup>63</sup>

2. Halo effect. No provision is made to guard against an extraordinarily high or low rating on some one trait affecting the other qualities and affecting the final score

---

59. Probst, J. B., op. cit., p. 16.

60. Mason, E. G., op. cit.

61. Walker, Lisle, in comments made regarding the Service Rating form on January 9, 1939.

62. Ibid.

63. Ibid.

in an unwarranted degree.<sup>64</sup>

3. Adjusting ratings. Adjustments must or should be made for the differences between high and low graders to insure uniformity in the rating.<sup>65</sup>

4. Tendency to average ratings. This is caused by requiring the rater to judge all items even though he lack sufficient information, the subsequent doubt as to where the rating should be,<sup>66</sup> and hesitancy to give unfavorable reports.<sup>67</sup>

5. Biased ratings. The tendency to either consciously or unconsciously rate a person high or low because of likes or dislikes are more apt to occur when ratings are made on such general terms.

6. Distribution of final scores too narrow. In use, the final scores tend to cluster around the upper half of the five classes instead of being distributed over the entire range as normally expected. Little differentiation is made between the high and low men; they are not segregated into the several classes. Two reasons for this are that only three items are rated and the lack of objectivity.

The more familiar rating scales have been covered.

---

64. Thorndike, E. L., "A Constant Error in Psychological Rating", Journal of Applied Psychology, Vol. IV: p. 25-29.

65. Scott, Walter Dill, Clothier, Robert C., and Mathewson, Stanley B., op. cit., p. 176.

66. Probst, J. B., op. cit., p. 17.

67. Brumbaugh, H. C., op. cit., p. 14.



There have been many other slightly different rating scales developed but they are probably some variation or combination of the ones already discussed.

Horizontal and Vertical Rating.

Any of the rating scales discussed could be given to the employees with instructions to rate all of the other employees on the scale. This mutual rating is described as horizontal. Vertical rating, which is the more common, occurs when supervisors rate the foresters on the Forest, and regional foresters rate members of their staffs and supervisors.<sup>68</sup>

---

68. Yoder, Dale, op. cit., p. 251-2.

**PART II**

**DEVELOPMENT OF A RATING SCALE**

## CHAPTER IX

## NEED OF A SPECIAL RATING SCALE FOR EACH GROUP

There is no one rating scale that is universally applicable for all types of work.<sup>1</sup> For example, the mental and physical makeup of a successful professional forester in the Federal Forest Service is quite different from that of an office clerk. Therefore, a rating scale for foresters should include and emphasize different traits than those included in a similar system for piece workers. "The soundest practice is for the management to develop its own rating scales, so that they may be well adapted...to the group...with whom they are to be used."<sup>2</sup>

The problem at hand is to devise a rating scale or service rating plan for professional foresters in the Federal Forest Service, particularly those engaged in supervisory and sub-administrative work, which gives as accurate a measure of operating efficiency as can be obtained, by considering the essential characteristics of the individual. Representative positions considered include junior forester, district ranger, and forest supervisor and staff, other than non-professional positions.

The development of a rating scale may be divided into

1. Burt, Harold Ernest, op. cit., p. 321; Yoder, Dale, op. cit., p. 264; and Walters, J. E., Applied Personnel Administration, p. 170.
2. Walters, J. E., op. cit., p. 170.

four well defined steps as follows:<sup>3</sup>

1. Obtain a list of the traits or qualities considered essential to the successful performance of a professional forester.

2. Tabulate the traits and select the ones most desired.

3. Devise a rating scale made up of the five to eight traits considered most desirable, and list the objective behavior under each trait.

4. Test the rating scale.

The development of the first three steps will now be described in detail while the fourth step will be described in a later chapter.

---

3. Ibid., p. 170-187; Yoder, Dale, op. cit., p. 264-6;  
Burt, Harold Ernest, op. cit., p. 320-329.

## **CHAPTER X**

### **SELECTING THE ESSENTIAL TRAITS OF A SUCCESSFUL FORESTER**

## CHAPTER X

## SELECTING THE ESSENTIAL TRAITS OF A SUCCESSFUL FORESTER

These essential traits were obtained through a questionnaire, sent to representative members of the profession, conference with members of the profession, and by consulting leaders who have written on the subject.

A copy of both the letter and questionnaire are shown in Figures 1 and 2.

---

---

---

Dear Sir:

I am writing a Master's thesis on rating scales in the graduate School of Forestry at Oregon State College.

In connection with this thesis it is necessary for me to get some information with reference to the personal qualifications of professional foresters in the Government service. Since this information should come from men who are active members, I would greatly appreciate your help on the questionnaire enclosed.

Very truly yours,

Figure 1. Letter Requesting Traits and Objectives  
for Rating Scale.

The questionnaire was sent to 12 members of the profession, selected to obtain representative opinions from different positions. These included three district rangers, one forest supervisor, one assistant forest supervisor, three men on regional forester's staffs, and four men in

the field of forest research. A sample questionnaire follows:

QUALITIES AND OBJECTIVE TRAITS DEEMED ESSENTIAL  
TO THE HIGHEST PERFORMANCE OF A PROFESSIONAL  
FORESTER IN THE FEDERAL FOREST SERVICE

(Please fill out and mail to Alva W. Blackerby,  
Oak Grove, Oregon)

What are the essential personal requisites of a professional forester in the Federal Service? (This might include intelligence, foresight, industry, honesty, health, etc.) Please list them in the spaces provided below and in the order of importance with the most important first.

After each essential trait, please list the objective behaviors by which you judge or rate each trait. (For example, intelligence might be judged by the rapidity with which one learns, or promptness in giving decisions, etc.; or personality might be rated by one's size and appearance, cheerfulness, sense of humor, etc.)

1. (Essential trait) \_\_\_\_\_
  - a. (Objective behavior) \_\_\_\_\_
  - b. \_\_\_\_\_
  - c. \_\_\_\_\_
  - d. \_\_\_\_\_
  - e. \_\_\_\_\_
2. (Essential trait) \_\_\_\_\_
  - a. (Objective behavior) \_\_\_\_\_
  - b. \_\_\_\_\_
  - c. \_\_\_\_\_
  - d. \_\_\_\_\_
  - e. \_\_\_\_\_

(And so forth up to and including 12 traits.)

(The decision as to the number of essential traits needed will be left to you. Leave some blank or add more, as you choose.)

Figure 2. Traits and Behaviors Questionnaire.

The questionnaire requested each man to list the essential

personal qualifications of a successful professional forester in rank order together with the objective behaviors by which these traits or qualifications are judged.

Answers were received from eight of the twelve men sent copies of the questionnaire. Two men, both district rangers, filled out the questionnaire as requested. One man gave some of his personal views in his letter and enclosed a copy of a lecture on this subject given by one of the regional foresters. Another man suggested source material where I might find this information. The other four men who answered were unable to be of assistance for various reasons.

The results of the questionnaire as shown by the two complete responses follows:

Essential Traits and Objective Behaviors by Which  
Each Trait is Judged

- |                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ol style="list-style-type: none"> <li>1. Character               <ol style="list-style-type: none"> <li>a. Honest</li> <li>b. Loyal</li> <li>c. Tolerant</li> <li>d. Sympathetic</li> <li>e. Firm</li> </ol> </li> <li>2. Personality               <ol style="list-style-type: none"> <li>a. Energetic</li> <li>b. Enthusiastic</li> <li>c. Virile</li> <li>d. Courtuous</li> <li>e. Courageous</li> <li>f. Habits and dress</li> </ol> </li> </ol> | <ol style="list-style-type: none"> <li>1. Ability to exercise good judgment.               <ol style="list-style-type: none"> <li>a. Judged by decisions made</li> <li>b. By analysis of their problems and worth</li> </ol> </li> <li>2. Resourcefulness               <ol style="list-style-type: none"> <li>a. Skill with which work is performed</li> <li>b. Means or ways of overcoming obstacles</li> </ol> </li> </ol> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|



## 3. Intelligence

- a. Anticipate behavior--understand human reaction.
- b. Analyze problems
- c. Solve problems
- d. Ability to learn and comprehend
- e. Ability to concentrate on essentials

## 4. Industry

- a. Accuracy
- b. Thoroughness
- c. Completeness
- d. Perseverance
- e. Speed

## 5. Adaptability

- a. Ability to meet strange situations
- b. Race and religious tolerance
- c. No caste distinctions--equality to mankind

## 6. Administrative ability

- a. Power of decision
- b. Value opinions of others
- c. Gives credit where due
- d. Discipline when necessary
- e. Be fair--don't have pets
- f. "Speak softly but carry a big stick"

## 3. Intelligence

- a. Ability to meet any situation
- b. Ability to use facts in such a way as to guide their action toward the desired results

## 4. Personality

- a. Inherent characteristics of individual

## 5. Perseverance

- a. Ability to overcome obstacles
- b. Persistence of each man toward his work, his manner of applying for a job, etc.

## 6. Adaptability

- a. Ability of each man to adjust himself to changing conditions

## 7. Initiative

- a. Ability of each man to think for himself
- b. Introduction of new ideas

## 8. Loyalty

- a. Ability to work for and with his immediate superior

9. Honesty
  - a. Power of each man to be straight forward in his work and relations with his fellow men
  - b. Ability of each man to take responsibility for his own actions
10. Enthusiasm
  - a. The interest displayed by men in or toward their work
  - b. Whether or not the man is "asleep on the job" or wide awake to better himself, his living and working conditions
11. Aggressive
  - a. Ability of each man to speak for himself.
  - b. Be able to tell what he is able to do
  - c. Is able to tell what he wants to do
  - d. His ability to stand up for his own rights
12. Ties in with personality but should be segregated to show ability of man to keep himself neat and clean around others
13. Physical fitness
14. Decision
  - a. Ability of the man to be decided
15. Tact
  - a. Diplomacy in dealing with others
16. Dependability
  - a. Can the man be relied upon? Ties in with honesty

## 17. Firmness

### a. Firm

Other members of the profession were consulted to determine their opinions as to the traits and qualities most important in a successful forester. These follow:

Illick, Professor of Forest Management at New York State College of Forestry at Syracuse University, writes, "The United States Forest Service in sizing up temporary employees for possible permanent employment stress the qualities of personality, reaction to criticism, physical strength, enthusiasm, honesty, courage, ability to work with others, and powers of observation".<sup>1</sup>

Moon and Brown, late Dean and Professor of Forest Utilization, respectively, at New York State College of Forestry at Syracuse University summarize the principal qualifications of a forester, as:

"1. A sound and technical training in a professional school of forestry.....

2. Resourcefulness.....Closely related to this feature is strong character, native honesty, keen sense of responsibility, and initiative.

3. Robust and vigorous health.

4. A spirit of public service and a strong desire to

---

1. Illick, Joseph S., An Outline of General Forestry, p. 288.

improve the conditions of the country with respect to its forest resources."<sup>2</sup>

Kylie, Hieronymus, and Hall stress, "A forester should be physically strong, an outdoor man, and willing to endure hardships. He should be intelligent, mentally resourceful, and congenial. He should be a nature lover, and should be willing to make personal sacrifices for his work."<sup>3</sup>

Graves, Dean of the School of Forestry at Yale University, and Guise, Assistant Professor of Forest Management at Cornell University, consider the important personal qualifications to be:<sup>4</sup>

First, "...a personal interest in the problems of forestry....It is the factor which enables men to meet with readiness the discomforts and even dangers of forest work and tenaciously holds them to it....A second qualification generally recognized as essential is adaptability....This refers to the ability to adjust oneself to new situations, environments and people....There are a number of traits that make for adaptability, an intellectual interest in unusual situations, an objective point of view in taking things as they come, keen perception of the background of given situations, and tolerance of others whose mode of life, point of view and manner of speech may be different from one's own."

Third, "The forester should have a high degree of self-reliance....He should have initiative, intellectual and otherwise. With this should go an active imagination. The forester is constantly meeting new problems, complex, and

2. Moon, Franklin, and Brown, Nelson Courtland, Elements of Forestry, p. 228.

3. Kylie, H. R., Hieronymus, G. H., and Hall, A. G., CCC Forestry, p. 292.

4. Graves, Henry S., and Guise, Cedric H., Forest Education, Chapters 4, 5, and 6.

difficult of solution.....He must often look far ahead, envisage possibilities he heretofore little recognized and be ready to initiate action that may not come to realization for a long period of time. Mental alertness and resourcefulness come into constant play in meeting emergencies and overcoming obstacles.

"In a great deal of the work in the general practice of forestry there is required executive ability, with all the mental and personality qualifications that are combined in this trait."

Mason, Consulting Forester, says that the qualities desired in a forester are character, intelligence, broad training, vision to see opportunities and courage to work for their realization.<sup>5</sup>

Kelley, Regional Forester, stresses the following qualifications for foresters: personality, physique, and interest, virility, intelligence, straight-thinking, quick-mindedness, mental balance, courage, self-starter, versatility, honesty, sincerity, temperance, interest in human welfare, yen for rendering service, love of the soil, reverence for things natural, never-satisfied inquisitiveness, keen powers of observation, sportsmanship, confidence with a freedom from egotism and complacency, tolerance, open-mindedness, specialized education in the basic sciences upon which forestry is built, and a good general education.<sup>6</sup>

Riebold divides the desirable qualifications of a

---

5. Mason, David T., "The Requirements for an Education of a Forester", Journal of Forestry, V. 35, p. 545-9.

6. See bottom page 84.

professional forester into two groups which are:

1. Ability to handle supervisory problems from "ability to understand and follow a plan, to foremanship and personnel managing ability; from ability to get results in a given time at a predetermined cost to maintaining harmonious working relations within a group of workers and between the group and himself; and the ability to direct, reprimand, praise, and train his crew members."

2. Aptitude in handling problems concerning "administrative or managerial ability, that is, the ability to plan, organize, staff, direct coordinate, budget, and report".<sup>7</sup>

The Forest Service states that the requisites to success include industry, honesty, soundness of character, a liking for the sort of life which he must lead, the health and constitution to stand the work, foresight, broad-mindedness, thoroughness in details, administrative and executive ability, teaching ability, ability to meet people and gain the confidence of the public, and a spirit of service.<sup>8</sup>

6. Kelley, Evan W., What the Forest Service Expects of Forest School Students, Lecture delivered at Joint Meeting of Northern Rocky Mountain Section, Society of American Foresters, with Conclave of Western Forest Schools, February 2, 1939.
7. Riebold, R. J., "In-Service Training for Foresters", Journal of Forestry. V. 36, p. 146-8.
8. Careers in Forestry, Misc. Publication No. 249, United States Department of Agriculture, January 1938, p. 4.

Assistant Forest Supervisor Elliott, in discussing the qualities in the order of their importance in the best type of forestry men, places common sense first. He says that it is something you can't get along without. In second place he puts technical ability and says that it is becoming more and more important. Physical ability to be on the job and handle the work takes third place, and in fourth place he puts personality to do public contact work.<sup>9</sup>

Kooh, Assistant Regional Forester, sheds some light on the relative importance of personal qualifications and technical ability when he says:<sup>10</sup>

"I have sat in on many conferences for selection of a man for promotion to a vacancy in the Forest Service, and almost invariably the selection of the man hinged upon the personal qualifications and experience of the candidate rather than the quality of his technical training. The two essentials are intelligence and character. By intelligence I mean the quality of the brain, the ability to learn, to reason, and to use judgment. Character is a broad term, but the essential qualities are confidence, courage, and leadership. Since much of the work of a forester is in the field, it is essential that the forester possess a certain ruggedness of spirit and body which will enable him to meet with equanimity and to enjoy the discomforts and occasional hardships of woods life, and to demand the respect of the tough men of the woods he must command and associate with....A forester who attains the highest positions in the profession must combine the ruggedness of the woodsman with the finer qualities and educated percep-

---

9. Elliott, Roy, from interviews with the writer during November, 1936.

10. Kooh, Elers, "Technical Requirements for a Forester in the Federal Service", Journal of Forestry, V. 35, p. 803-9.

tions of the man who is at ease in the councils of the learned as well as in the logging camp."

The School for Forestry at Oregon State College has a particular type of man in mind when they give the qualifications of a forester. They say:<sup>11</sup>

#### "THE QUALIFICATIONS OF A FORESTER

In common with every other profession, forestry has its standards, its requirements, and its ideals. In the quarter century which has elapsed since the appearance of the American forester, certain well-defined things have come to be recognized as requisite in every true member of the profession. No one of these attributes can be regarded as of greater importance than the others. In technical training the forester must measure up to the standard set by the best men in the profession. He must have that peculiar type of honesty which makes him demand of himself a full measure of service even though for days and weeks he may be out of touch with his superior officer. He must be loyal to his profession and to his fellows, to those who give him orders, and to those whom he directs. He must have initiative, for his work frequently places him in a situation where he must make his own decisions and formulate his own plans. He must school himself in teamwork, for only through whole-hearted cooperation can he become of greatest service in his profession. He must have vision and with his vision, faith, for the real fruits of his labor may not mature until years after he has made his exit from the stage of life. He should unselfishly strive for the betterment of his profession. His criticisms should be constructive. His judgments should be withheld until he has full possession of all the facts. He should recognize his obligation to his school which trained him and to his state which made this training possible. He should be true to his ideals. With all this he may then be a forester."

The inclusive quality of leadership<sup>12</sup> has been named

---

11. Oregon State College, School of Forestry, Monthly Bulletin, issued March, 1924.

12. See bottom page 87.



by some as being essential. Others have included, in their listing of prerequisites, basic qualities which go to make up leadership. Desiring additional information on this point the following question was asked several professional foresters: "Is leadership desirable or even essential in a forester? Please give reasons for your answer. (Leadership is defined by Tead<sup>13</sup> as '...the activity of influencing people to cooperate toward some goal which they come to find desirable'.)"<sup>14</sup> All of the answers were in the affirmative. Some of them are as follows:

Steele, Assistant Forest Supervisor of the Mt. Hood National Forest, stresses the need of leadership in public service when he says:<sup>15</sup>

"For a forester in public service, leadership is most desirable and in most cases essential. Desirable forest practices on public lands can only be accomplished with public support and the support of public opinion can only be gotten through the leadership of the forester. Another reason why leadership is essential is that forestry in the United States is new. The public as a whole knows practically nothing about the needs of forestry--it couldn't, as a whole, distinguish between good and bad practices.

---

12. Leadership will be defined as "...the activity of influencing people to cooperate toward some goal which they come to find desirable".<sup>13</sup> With reference to foresters this includes activities involving the general public to follow the policies deemed best for the common good as well as activities with subordinates.

13. Tead, Ordway, The Art of Leadership, p. 20.

14. Ibid.

15. Steele, Foster, from an interview with the writer on March 4, 1939.

Looking back we see many men in forestry who are cogs in a wheel, not assisting as far as the public is concerned in shaping opinion or in educating the citizenry. Only through this leadership can we hope to accomplish what must be accomplished in the United States and the more district rangers, guards, etc. that we can get with proper perspective and leadership the more progress we can make."

Brown, District Ranger at Estacada, Oregon, stressed the need for leadership when dealing with subordinates in the following words:<sup>16</sup>

"Leadership is absolutely required in a forester because foresters going into administrative work will have others working for them. To secure results the forester or district ranger must be a leader. He must guide the thoughts of his subordinates. He should be able to display the following qualities: will power, knowledge, enthusiasm, self confidence, energy, responsibility, and courage."

Thompson, Assistant Regional Forester of the Division of Personnel Management at Missoula, Montana, points out that the lack of leadership is the major cause of most personnel cases. Mr. Thompson writes:<sup>17</sup>

"I would like to emphasize that I think 'Leadership' is not only desirable but absolutely essential in any forester who hopes to climb the ladder of success in the administrative grades of ranger, supervisor, regional forester, etc. In fact the greater number of our personnel cases arise among men who lack this very essential quality but who may otherwise be intelligent, of good character and of high scholastic attainment."

- 
16. Brown, Carlos T., taken from a letter to the writer, dated March 9, 1939.  
17. Thompson, P. A., taken from a letter to the writer, dated March 6, 1939.

The writer also desired to have some definite information as to whether the quality of leadership is desirable in professional foresters in the field of research as well as in administrative and supervisory work. Therefore the following question was asked some of the same professional foresters. "Is leadership an important accomplishment for a professional forester in the field of research? Why or why not?" Representative replies include:

"Yes. Research men work with others. They must guide the thoughts of the general practitioner in the field in order to get them to fall into line with the new facts, methods and practices found in research. If this were not done research would not amount to much."<sup>18</sup>

"Yes. Too many brilliant minds are obscured because of the lack of advertising. To sell an idea or fact, proven through research, a person must possess qualities of leadership. Any man in any field of endeavor should strive to acquire and improve qualities, habits, mannerisms, knowledge, etc. which distinguish the leader."<sup>19</sup>

Thompson puts it this way:<sup>20</sup>

"This quality of 'Leadership' may not be so important for a professional forester in the field of research but nevertheless if he progresses very far in research he should be able to direct the work of subordinates and exert a commanding influence over individuals and the public in making application of the results of his research work."

After establishing the fact that leadership is vital

---

18. Brown, Carlos T., op. cit.

19. Aydelott, Owen L., from a letter to the writer received March 11, 1939.

20. Thompson, P. A., op. cit.

in a progressive professional forester and noting that it is a broad quality, dependent upon several basic qualities, it was decided that some effort should be made to discover what fundamental qualities are essential to leadership.

Several authorities were consulted. Some findings follow:

According to Wohlenberg leadership involves courage, intelligence, initiative, and knowledge. "The two fundamentals of leadership which have been working overtime are knowledge and a sort of latent intelligence. The two fundamentals which have been practically dormant are courage and initiative."<sup>21</sup>

Schell, Professor of Business Management, Massachusetts Institute of Technology, says, "These three traits; interest in and an affection for people, power of personality, and a scientific trend of mind, may be said to be the outstanding requirements of executive success". He adds that executives should also possess moral uprightness, physical vigor, and intelligence but points out that these qualities are needed wherever accomplishment is the goal."<sup>22</sup>

Craig, Assistant Professor of Industry, and Charters, Director, Research Bureau for Retail Training, and Dean of the Graduate School at the University of Pittsburg define the qualities of personal leadership as forcefulness, abil-

---

21. Wohlenberg, E. T. F., "Leadership in Forestry and Lumbering", Journal of Forestry V. 31, p. 308-10.

22. Schell, Erwin Haskell, The Technique of Executive Control, p. 12-13.

ity to take a personal interest in the men, ability to get the work done correctly, ability to get and use the ideas of men, ability to be one of the men, ability to lead rather than boss the men, ability to develop teamwork, ability to show kindness without being considered easy, ability to reprimand properly, ability to keep from worrying, ability to delegate work properly, ability to call forth the best efforts of the men, ability to train men on the job, ability to make a new man feel at home, and self-confidence.<sup>23</sup>

Tead, who is a Lecturer in Personnel Administration at Columbia University, says, "Qualities necessary in leaders which seem ideally desirable are: Physical and nervous energy, a sense of purpose and direction, enthusiasm, friendliness and affection, integrity, technical mastery, decisiveness, intelligence, teaching skill, and faith."<sup>24</sup>

Griffing, Educational Advisor of the Ninth Corps Area of the United States Army, groups these qualities of leadership under five heads: Personality, intelligence, skill and knowledge, energy and initiative, and character.<sup>25</sup>

23. Craig, David R., and Charters, W. W., Personal Leadership in Industry, p. 235-6.

24. Tead, Ordway, op. cit., p. 83.

25. Griffing, John B., Fundamentals of Leadership, (Headquarters Ninth Corps Area, Office of the Educational Adviser, Presidio of San Francisco), Lesson 2, p. 1.

CHAPTER XI

TABULATION OF THE TRAITS

## CHAPTER XI

### TABULATION OF THE TRAITS

Upwards of a hundred different terms are used, by the foresters quoted, in describing the essential qualities of a professional forester. When these terms are closely analyzed it is found that a great many are merely synonymous, or descriptive parts of each other. They were classified, first, into two groups consisting of fundamental or singular qualities and plural qualities. The plural qualities are a result of the combination of two or more basic qualities. For example, "intelligence" represents, largely, a basic and constant quality while "specialized forest education" represents and is dependent not only upon intelligence but physical and nervous energy, character, and an opportunity for their interplay as well. All of these basic qualities have to be synchronized and harmonized before a specialized forest education results.

The importance of recognizing these plural qualities agrees with the Gestalt Psychologists who propagate the idea "...that the whole is something quite different than its summated elements..."<sup>1</sup> They emphasize the fact that activities such as learning and application of executive ability are not the simple sum of their basic elements but

---

1. Trow, William Clark, "How Shall Teaching Be Evaluated?" Educational Administration and Supervision 20:264-72.

a synchronized total pattern of all the basic parts in their environment. This plural quality is something more than the sum of its parts just as a football team is much more than eleven non-synchronized athletes. Trow illustrates this principle as it applies in rating teaching ability when he says, "If, for the sake of the game, one wishes to use the sum of numerically evaluated traits as the score, and the person who gets the highest score wins, all well and good; but these scores, even granting that the traits are accurately judged, should not be confused with teaching ability". He adds that it is the total pattern of the various characteristics that counts, and how they harmonize in their social setting.<sup>2</sup> This emphasizes the importance of judging a man by his performance and success in his work.

The basic qualities and the plural qualities were next divided into four and two respective divisions of nearly synonymous characteristics. The outstanding term or combination in each group was used as the name of the group. The essential qualities of a forester, as given by the thirteen references consulted were tabulated to see how many times each of the six groups were selected as essential. Not more than one vote was given any one reference on any one of the six groups even though a similar

---

2. Ibid.



quality was named in more than one way. For example, one man listed honesty and sincerity but they were both counted as one vote under the quality of character. The titles of the four groups of basic qualities and the two groups of plural qualities together with the number of times that each group was selected as essential follow:

Basic qualities:

1. Intelligence-- listed by 8 references as essential.
2. Personality-- listed by 6 references as essential.
3. Character-- listed by 11 references as essential.
4. Physical and nervous energy--listed by 11 references as essential.

Plural qualities:

1. Scientific trend of mind or common sense-- listed by 9 references as essential.
2. Skill and knowledge-- listed by 10 references as essential.

In addition, leadership is mentioned several times. Inasmuch as it has been established that leadership is essential in a successful forester and because the type of leadership needed in a forester is dependent upon several qualities, the opinions of the five references, who are reporting the essential qualities in leadership, were also tabulated.

The tabulated opinions of the eighteen references, thirteen listing essential qualities in foresters and five

listing essential qualities in leaders, follows:

Basic qualities:

1. Intelligence-- listed by 13 references.
2. Personality-- listed by 10 references.
3. Character-- listed by 16 references.
4. Physical and nervous energy-- listed by 16 references.

Plural qualities:

1. Scientific trend of mind or common sense-- listed by 12 references.
2. Skill and knowledge-- listed by 14 references.

For clarification each of these six titles, in the two groups, are briefly defined as follows:

Basic qualities:

1. Intelligence--Intellectual capacity which determines how well we do things and demonstrated by our ability to meet new situations.<sup>3</sup>
2. Personality--The way an individual impresses other folks.<sup>4</sup>
3. Character--What a man is on the inside.<sup>5</sup>
4. Physical and nervous energy--Drive, endurance, and vigor of mind and body.

Plural qualities:

1. Scientific trend of mind or common sense--Ability to

---

3. Chambers, O. R., from notes taken in a psychology class, May 4, 1938.

4. Griffing, J. B., op. cit., Lesson 2, p. 2.

5. Ibid., Lesson 3, p. 1.

solve problems impartially and wisely in the light of all available evidence.

2. Skill and knowledge--Ability to do things resulting from training, education, and experience.<sup>6</sup>

It has already been pointed out that a great many essential qualities given by the references are very similar or descriptive parts of each other. The qualities named by the references that were considered and counted under each group are in the following list. Some items appear to be borderland cases and it is difficult to limit them to one definite group.

1. Intelligence:

- Quick-mindedness
- Mental alertness
- Inquisitiveness
- Ability to learn and to reason
- Ability to meet new situations
- Adaptability
- Active imagination
- Resourcefulness
- Capacity to see the point
- Capacity to sense relationships
- Capacity to put "two and two" together
- Ability to recognize salient points

2. Personality

- Habits
- Dress
- Tact
- Congeniality
- Ability to meet people
- Personal appearance
- Enthusiasm
- Energetic attitude
- Friendliness and affection
- Cheerfulness

---

6. Ibid., Lesson 2, p. 4.

Poise

### 3. Character

Temperance  
Sincerity  
Honesty  
Loyalty  
Tolerance  
Sympathy  
Courtesy  
Firmness  
Broadmindedness  
Dependability  
Faith  
Courage  
Self-reliance  
Sportsmanship  
Teamwork  
Cooperative  
Ability to work with others  
Spirit of public service  
Love of nature and things natural  
Interest in human welfare  
Interest in rendering service  
Reaction to criticism  
Willingness to make personal sacrifices  
Integrity  
Reliability

### 4. Physical and nervous energy

Industry  
Initiative  
Self-starter  
Perseverance  
Aggressiveness  
Physically strong  
Certain ruggedness of spirit and body  
Willingness to endure hardships  
Outdoor man  
Virility  
Drive  
Endurance  
Vigor of body and mind  
Quality that begets zeal and enthusiasm  
Will to win  
Dynamic emotion  
Abounding energy  
Robust

## 5. Scientific trend of mind or common sense

- Use of good judgment
- Vision
- Foresight
- Mental balance
- Straight thinking
- Open-mindedness
- Decision
- Executive ability
- Firmness
- Confidence
- Ability to reprimand
- Ability to direct
- Ability to plan
- Ability to coordinate
- Sense of purpose and direction

## 6. Skill and knowledge

- Technical training
- Knowledge
- Specialized education
- Broad training
- Powers of observation developed
- Versatility
- Ability to train others
- Ability to understand and follow a plan
- Thoroughness in details
- Personal interest developed in problems of forestry
- Technical mastery
- Technology and processes by which purposes are realized

## CHAPTER XII

### DEVISING THE RATING SCALE

## CHAPTER XII

## DEVISING THE RATING SCALE

After the important qualities have been determined the actual rating scale must be developed. Several principles which aim to increase the efficiency of rating scales and their methods of use have been formulated that may serve as guides in the actual construction, subsequent development, and administration. The principles will now be considered under two main heads entitled features of sound rating scales and approved administrative features.

Features of Sound Rating Scales:

1. Make all phases of the rating system as objective as it is possible to make them in order to avoid ambiguity and personal interpretations.<sup>1</sup> Some specific methods of avoiding this danger will be discussed in other points.

2. Determine traits or items selected for rating purposes in the light of the work actually performed by the employees.<sup>2</sup> If a man isn't required to do public speaking in his work he shouldn't be rated upon his ability as a public speaker. One man says, "Efficiency rating systems that get away from an analysis of the duties, get away from

---

1. See Mosher, William E., and Kingsley, J. Donald, op. cit., p. 441; Viteles, Morris S., op. cit., p. 212; Burt, Harold Ernest, op. cit., p. 326-8.

2. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 441.

the facts and the further they get away from the facts the less their value".<sup>3</sup>

3. Refer to but one type of activity carried on, or to but one type of result achieved by the men to be rated in each quality.<sup>4</sup> Caution should be exercised to see that the quality is not a composite of several abilities that vary independently.<sup>5</sup>

4. Mosher and Kingsley state, "Rating scales should be devised in such a way that the points in the continuum from the least to the highest are equi-distant from one another".<sup>6</sup> In other words, the traits should be so arranged and stated with reference to each other that the whole list of traits will divide the total range into spaces of approximately equal length.

5. Weighing of the essential qualities in the scale should be based upon relative importance of these qualities. This may be done by a conference or some other consensus of opinion secured from leaders in that field. Occasionally it is done experimentally by comparing each trait with a criterion, when such a criterion is available. Usually it is determined rather arbitrarily by using the best judgment of those familiar with rating scales and the

---

3. Faine, P. L., excerpt sent to the writer, Feb. 22, 1939.

4. Viteles, Morris S., op. cit., p. 212.

5. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 136.

6. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 442.



occupation.<sup>7</sup>

6. State the qualities or traits in the every day thinking of the rating officer.<sup>8</sup> Supervisors and not psychologists are going to rate the employees. Therefore, assistance should be rendered through the use of simple questions, words, and examples, that bear directly on the work that the men are doing. However, this can be carried to extremes.<sup>9</sup>

7. Define qualities specifically and in some detail. When they are named only, different interpretations result, and ratings tend to be inconsistent from one period to another and from one rater to another. For instance if the scale mentioned cooperation, the rater might think of it as meaning either the ability to get along with co-workers or willingness to carry out orders.<sup>10</sup>

8. Qualities may, in some cases, be described but not named, so that raters may avoid personal connotations that are at variance with the characteristic as it is defined in the rating scale or by another rater.<sup>11</sup>

9. State the items to be checked in concrete terms

- 
7. See Burt, Harold E., op. cit., p. 323-7; Yoder, Dale, op. cit., p. 266-7, 261.  
 8. Probst, J. B., op. cit., p. 22.  
 9. See Baridon, Felix E., and Loomis, Earl H., op. cit., p. 195; Kingsbury, Forrest A., "Analyzing Ratings and Training Raters", Journal of Personnel Research, V. 1, p. 377-383.  
 10. See Yoder, Dale, op. cit., p. 255-6; Burt, Harold E., op. cit., p. 326.  
 11. Yoder, Dale, op. cit., p. 265.

and these items should be concerned with observable behavior,<sup>12</sup> such as may be noted during the regular course of the work. In place of asking for a rating on such a trait as initiative, ask to what extent the ratee starts things, plans activities, and starts others. The rater merely checks the behavior he believes to be associated with the ratee. One form requires the rater to substantiate the score by referring to some illustrative incident.<sup>13</sup> The more concretely the ability is expressed the greater is the expectation that the various raters will be judging the same thing. "Such traits and characteristics should be significant for the successful performance of the work assignments. ...."<sup>14</sup>

10. Include only qualities which are non-measurable by exact means. If typing speed is deemed important, a record may be kept in the personnel report but the quality should be measured with accurate instruments.<sup>15</sup>

11. Include only qualities that appear in varying degrees among those to be rated. An example is sex, and some suggest that honesty and loyalty should also be included under this heading.<sup>16</sup>

---

12. Laird, Donald A., op. cit., p. 191.

13. Yoder, Dale, op. cit., p. 255.

14. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 441-2.

15. Yoder, Dale, op. cit., p. 264.

16. Ibid.

Mosher, William E., and Kingsley, J. Donald, op. cit., p. 442.

12. Avoid such general terms as good, fair, medium, or excellent. Such terms are not consistently used by different raters nor by the same rater at different times.<sup>17</sup> Such appraisal steps should be specifically defined in order to overcome this tendency of raters to vary their quantitative appraisals of various qualities.<sup>18</sup> The rating steps developed in this study are "almost always", "usually", "about half of the time", "occasionally", and "hardly ever". They are used as positive answers to statements which are, in effect, questions, such as, "He appears cheerful and cordial".

Sometimes several appraisal steps are defined at length, similar to the Probst Service Scale, and the rater checks the correct step. Thus instead of simply asking raters to rate personal appearance they might be asked to begin by answering the question, "Does he favorably impress people by his size, bearing and manner?"

13. Avoid using extreme descriptive phrases that will be evaded. For example, in rating an employee on promptness such extreme descriptive terms as "always prompt" and "never prompt" should be avoided. The result if a scale was constructed in this way, and had five descriptive steps, would be to limit the range to the middle three steps. The de-

---

17. Yoder, Dale, op. cit., p. 266.

18. Ibid., p. 255.

scriptive phrases should be arranged so that they will conform to a normal distribution curve.<sup>19</sup> Bingham and Freyd<sup>20</sup> say that if there are five phrases, the intermediate ones should be closer in meaning to the central one than to the extremes in order to spread the distribution.

14. Use three to five descriptive phrases in rating each trait. More than five divisions are inadvisable, because most raters will be unable to distinguish more than that many degrees of quality nor will they be able to grasp the range of the trait as a unit. At least three are necessary to provide a serviceable range of discrimination.<sup>21</sup>

15. Direct quantitative terms should be avoided. The immediate use of percentages, other numerical scores, or such words as high, superior, and low by a rater in classifying a ratee on some quality is undesirable. They give a sham impression of accuracy and fineness, but, "since they are not defined and in most instances are undefinable, no good can come from their use".<sup>22</sup>

16. Some men advise shifting the favorable extremes of a multiple-step or adjective type rating scale haphazard-

---

19. Laird, Donald A., op. cit., p. 192-3; and Yoder, Dale, op. cit., p. 266.

20. Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 137.

21. Ibid., p. 136; Yoder, Dale, op. cit., p. 266; and Laird, Donald A., op. cit., p. 191.

22. Laird, Donald A., op. cit., p. 191.

ly to the right and to the left,<sup>23</sup> while others oppose it.<sup>24</sup> Proponents say it reduces halo effects, while opponents say it is time consuming and irritating.

17. Recognize the natural reluctance of raters to rate negative qualities and their desire to say good things about a rater.<sup>25</sup> Some of this reluctance can be eliminated by describing the employee with observable facts and avoiding words that suggest a feeling of failure due to some past association. For example, such a feeling might be present if any of the letters "D", "E", or "U" were used in describing a man because of a feeling of antipathy toward them which had been developed in school.

18. Provide an optional selection of traits or qualities for the reporting officer, so that he may report on only those things with respect to which he has definite knowledge.<sup>26</sup> Yoder says, "If the rater feels uncertain with respect to any quality, he should not rate on that one".<sup>27</sup>

19. Directions to the rater should be clear, concise, and unambiguous. In addition to clearly explaining the mechanics of the rating process, they should warn the rater against the most common shortcomings,<sup>28</sup> and the directions

---

23. Ibid., p. 194; and Bingham, Walter Van Dyke, and Freyd, Max, op. cit., p. 138.

24. Probst, J. B., op. cit., p. 19.

25. Ibid., p. 22.

26. Probst, J. B., op. cit., p. 22.

27. Yoder, Dale, op. cit., p. 275.

28. Ibid.

should emphasize that judgments should be based upon performance on the job<sup>29</sup> to which the employee is assigned. Failure of raters to follow this principle results in comments such as these of Paine's who says, "I think, we, as raters, are prone to forget what the employee is supposed to do, as outlined in his job description, and we rate by comparison with another employee who may be doing different work".<sup>30</sup>

20. Use a rating scheme that people like. It is obviously desirable if a scheme is interesting, quickly and easily completed, and gains the interested cooperation and enthusiasm of the raters using it. Therefore, for the sake of time-saving and the securing of intelligent cooperation the number of items should be reduced to a minimum and made as interest-appealing as possible.<sup>31</sup>

21. Rate all employees on a single trait at a time. That is, rate all ratees on the first quality, then rearranging the order at random, rate them on the second quality and so on for the remaining qualities. Laird says, "This precaution avoids a shifting standard. Since all are rated on the same trait at the same time, the standard pre-

---

29. Baridon, Felix E., and Loomis, Earl H., op. cit., p. 197.

30. Paine, P. L., taken from a letter to the writer, dated February 22, 1939.

31. Laird, Donald A., op. cit., p. 193-4; and Mosher, William E., and Kingsley, J. Donald, op. cit., p. 441.

sumably remains the same".<sup>32</sup>

22. Sufficient understandable information should be provided on the rating scale for the ratee to make the proper adjustments when he sees his rating. It is very difficult for a man to see himself as others see him. So, if an employee finds out his strong and weak points, something he often doesn't know, he will know where to begin improving himself.<sup>33</sup> In this way the employees may also in a general way determine the fairness of the ratings.<sup>34</sup>

23. Convert total numerical ratings into final ratings consisting of approximately five classes. These may be labeled A, B, C, D, and E or entitled excellent, good, average, fair, and poor, etc. The basis for this type of grouping is the presumption that the traits rated yield a nearly normal distribution, hence the total ratings will yield a nearly normal distribution.<sup>35</sup> Two methods are commonly used in determining the limiting points between the groups. All of the total scores are arranged in a frequency distribution from high to low. The distribution is then divided into five groups with the highest ten per cent in the highest group or given an A rating, the next twenty per cent are given the B rating, the next forty per cent are given a C

---

32. Laird, Donald A., op. cit., p. 191; see also, Viteles, Morris S., op. cit., p. 212.

33. Walters, J. C., op. cit., p. 174.

34. Probst, J. C., op. cit., p. 22.

35. Burt, Harold Ernest, op. cit., p. 345.

rating, the next twenty per cent are given a D rating, and the lowest ten per cent are given an E rating. The dividing lines mark the limiting points between the groups. A key to the final ratings can then be prepared and used as a future key for converting total scores into final ratings.<sup>36</sup> The other common but more refined method is done by dividing the total frequency distribution into five groups in terms of standard deviations from the mean or average of the group. The number of standard deviations to apply for each class are found by consulting a table of normal expectancies<sup>37</sup> in per cents and finding the number of standard deviations for the above named per cents or any others decided upon.<sup>38</sup> Similar methods are used in correcting high and low rating tendencies in raters.<sup>39</sup>

24. Provide space for noting outstanding proficiencies and deficiencies, particularly those justifying promotion, demotion, transfer, etc.<sup>40</sup> An average rating may be highly deceptive in picking a man for a promotion because a total score often fails to properly harmonize a man's possible shortcomings and qualifications. Of course, the rating report should be studied in detail, too, in relation to the

---

36. Patterson, Donald G., op. cit., p. 365.

37. Garrett, Henry E., op. cit., p. 110.

38. Clinton, R. J., from notes taken in a class in Educational Statistics, December 11, 1938.

39. Burt, Harold Ernest, op. cit., p. 156-169 and p. 345-348.

40. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 442.



qualities that are considered most important in the particular situation at hand.<sup>41</sup>

25. Create more than one type scale when widely different occupations are rated. "Just as different types of positions require distinctive abilities, so the rating of those who function in such positions require different scales."<sup>42</sup> Scales for some positions will emphasize leadership, tact, originality, etc., while others emphasize co-operation, skill, industry, etc.

26. Continue research, analysis, and adjustment after the rating scale has been constructed in order to improve it and to keep abreast of changes. A change might be needed due to a shift in the functions of the position, which would change the relative importance of the essential qualities to be rated.<sup>43</sup>

#### Approved Administrative Features.

1. The central personnel agency should set up the rating scales with the cooperation of the supervisory and executive staff. Scoring should be handled in the personnel division.<sup>44</sup>

2. Accumulate and file the records concerning supervisor's estimates of employees in advance of any emergency

---

41. Laird, Donald A., op. cit., p. 194.

42. Yoder, Dale, op. cit., p. 264.

43. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 442.

44. Ibid.

requiring such information as a basis of decision on promotion, demotion, transfers, etc.<sup>45</sup>

3. Call for ratings not more often than twice a year.<sup>46</sup>

4. Don't expect average scores to tell the whole story. A high rating in one quality such as physical and nervous energy plus a low rating in intelligence gives an average numerical score but it does not necessarily mean that the employee's value to the service is average. Average scores are convenient but highly deceptive in actual use. For instance, if you are looking for an employee with an outstanding speaking voice or a fast worker, he can be found by carefully studying the applicable parts of the report but not by surveying the average scores.<sup>47</sup>

5. Keep a record of attendance, tardiness, promotion history, record of errors, and other personal records in the employee's personnel record. This facilitates its use in connection with promotions, demotions, etc. However, attendance and tardiness should not be rated as are the traits and characteristics.<sup>48</sup>

6. Instruct and train rating officials carefully in the use of rating scales. Teach them to carefully observe

---

45. Viteles, Morris S., op. cit., p. 212.

46. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 442.

47. Laird, Donald A., op. cit., p. 194.

48. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 442.

and analyze the significant traits in each employee, and to report them with accuracy and care on the rating scale. A raters' manual can be prepared which describes in detail the scale and proper methods of using it. It might well discuss the various personal qualities, their significance, ways in which they are manifested, and methods of increasing one's reliability in judging them. Typical errors and suggested ways of guarding against them could be enumerated and discussed. The raters should also be stimulated to give careful, conscientious, and intelligent consideration to their rating responsibility.<sup>49</sup>

7. Raters should not rate an employee unless he has a knowledge of him. A common difficulty which results arises out of the fact that the rater does not know the ratees well but hesitates to admit the fact. It is better that no ratings be made in such cases.<sup>50</sup>

When a man in the Forest Service is rated on a service of less than ninety days the notation is made after his name, "Less than 90 days".<sup>51</sup>

8. Have as many independent ratings on a person as conditions permit, as long as all of the raters are well

---

49. Ibid., and Kingsbury, Forrest A., "Analyzing Ratings and Training Raters", *Journal of Personnel Research*, V. 1. p. 377-383.

50. Yoder, Dale, op. cit., pp. 263-4 and 275; and Laird, Donald A., op. cit., p. 193.

51. United States Civil Service Commission, Preparation of Efficiency Ratings, p. 4.

enough acquainted with the work of those being rated as to be competent in their judgments. These raters should work independently. These separate ratings should be combined, and the average found, which should be used as the index or official rating for that person. Inaccurate ratings may be detected which may be caused if any of the raters are lacking in a knowledge of an employee's real ability. Radical disagreements may be due to rater bias or prejudice. In any case it affords an opportunity for a fruitful conference of the raters.<sup>52</sup>

Kornhauser, after testing a number of ratings made at the University of Chicago, concluded that the reliability of average ratings increases steadily up to four and that beyond four separate ratings entering into the average there is no change in reliability.<sup>53</sup>

9. Correct statistically any tendency of a rater to rate too high or too low. Frequently, two raters will agree, in general, as to the type of person the ratee is, and the general pattern of the two rating will correspond, but one will be generally lower throughout all traits than

---

52. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 439; Laird, Donald A., op. cit., p. 193; Viteles, Morris, op. cit., p. 212; and Paterson, Donald G., op. cit., p. 375.

53. Kornhauser, W. W., "What Are Rating Scales Good For?" *The Journal of Personnel Research*, V. 5, p. 189-193. See also, Rugg, Harold, "Is the Rating of Human Character Practical?" *Journal of Educational Psychology*, V. 13, p. 63.

the other. This tendency for one rater to judge his employees too leniently, thus giving all of them very high rating while another rater judges his employees too severely, thus giving all of them low ratings is described as the systematic error. One method of getting a measurement of this systematic error in scoring is to compare average ratings.<sup>54</sup> If one rater's ratings consistently average twenty points higher than another rater's ratings on similar groups of men, it may be assumed that the first rater tends to score much higher than the second rater. It is frequently possible, by careful analysis, to calculate weights to be applied to each rater's scores whereby the work of the various raters is made comparable. Efforts should be made to reform raters who tend to rate too high, too low, or who tend to rate all of their employees the same. Differences between their rating and those of other supervisors can be pointed out but it is usually more satisfactory to statistically adjust total ratings for high and low tendencies.<sup>55</sup>

10. Encourage employees to inquire as to their standing in order that ratings may be made as fruitful as pos-

---

54. For additional information see Yoder, Dale, op. cit., p. 267-70; and Burt, Harold E., op. cit., pp. 345-348.

55. Yoder, Dale, op. cit., p. 267-70; Paterson, Donald G., op. cit., p. 375; Viteles, Morris, S., op. cit., p. 212; and Laird, Donald A., op. cit., p. 194.

ible. Those falling below par should be informed of this fact and the reasons for it. Remedial work or some other adjustment should be made. If this part of the rating program is wisely handled it may become a worth-while instrument for improving standards.<sup>56</sup>

11. "Opportunity for appeal should be afforded as in connection with other decisions affecting the interests of the workers."<sup>57</sup>

12. Afford opportunities for and encourage employees to cooperate in setting up and revising scales, and passing on scoring methods. By publicising the whole rating program each employee may come to understand the plan, become convinced of its fairness, and learn on what types of behavior emphasis is being placed. A handbook is suggested as a good means of publicising the scheme.<sup>58</sup>

As far as was possible the principles of sound rating scales, just discussed, were incorporated into the rating scale constructed. In addition, the approved administrative features were kept in mind in an attempt to make it usable as well as a valuable tool for management.

The actual rating scale construction, with the six qualities selected as most essential in a professional forester in the Forest Service as a basis, will now proceed.

---

56. Mosher, William E., and Kingsley, J. Donald, op. cit., p. 442-3.

57. Ibid., p. 443.

58. Ibid., p. 442.

### Weighing the Essential Qualities.

Ten professional foresters in the field of forest education were given a list of the six qualities considered essential together with their definitions. They were asked to carefully consider the importance of these qualities in a professional forester in the Federal Forest Service and to list them in rank order with the most important first. The results of this poll are shown below in Table I.

TABLE I  
Evaluation by Rank Order of Six Essential Qualities  
in Foresters

Qualities	Rank order selection by 10 foresters										Rank order according to average
	1	2	3	4	5	6	7	8	9	10	
1. Intelligence	1	1	5	2	1	3	2	4	3	4	3
2. Personality	3	3	1	3	3	1	4	1	1	3	1
3. Character	2	2	2	1	2	2	3	2	5	2	1
4. Physical and nervous energy	4	5	4	6	6	4	6	3	2	5	5
5. Scientific trend of mind and common sense	5	4	6	4	4	5	1	5	4	1	4
6. Skill and knowledge	6	6	3	5	5	6	5	6	6	6	6

An examination of the results discloses varied opinions. Personality and character were selected by the group

as a whole as equally important, although personality was selected as low as fifth choice by one man. Intelligence was placed third by the group as a whole and almost on a parity with the first two. Physical and nervous energy and skill and knowledge were considered least important. However, one man selected physical and nervous energy as second choice and another ranked skill and knowledge in third place.

Correlations<sup>59</sup> were made between the ranking made by the several men in order to further study the relationship between the different rankings. The following coefficients of correlations were found:

---

59. The correlation coefficient tells the degree of similarity that exists between any two series and is expressed by means of an index figure ranging from plus 1.00 through .00 to minus 1.00. Thus, a coefficient of plus 1.00 indicates that the two series or rankings are identical, while a coefficient of minus 1.00 indicates that the two series are completely reversed, the highest rank in one being the lowest rank in the other. A coefficient of .00 represents a purely chance order of arrangement. A general guide is that a coefficient of from .00 to plus or minus .20 denotes indifferent or negligible relationship; plus or minus .20 to plus or minus .40 denotes low correlation, present but slight; plus or minus .40 to plus or minus .70 denotes substantial or marked relationship; and plus or minus .70 to plus or minus 1.00 denotes high to very high relation. Correlations shown throughout this study have been determined by the Spearman Rank-Difference method unless otherwise stated. For further information see Garrett, Henry E., op. cit., pp. 251-408.



Between judges number 1 and number 2	--	.94
Between judges number 3 and number 4	--	.26
Between judges number 5 and number 6	--	.57
Between judges number 7 and number 8	--	.14
Between judges number 9 and number 10	--	.09
Average <sup>60</sup> coefficient of the group	--	.50

This average correlation coefficient of plus .50 indicates that there is a definite positive relationship. However, interpreted in another way, it says that only thirteen times out of a hundred are the relationships closer than pure chance,<sup>61</sup> or what one would expect by drawings from a hat.

The results as a whole are interpreted to mean that all six qualities are vital but physical and nervous energy and skill and knowledge are considered of somewhat less importance than the other four qualities. The weights given the six qualities on the basis of a total of 1008 points follow:

1. Intelligence	180 points
2. Personality	180 points
3. Character	180 points
4. Scientific trend of mind and common	

---

60. The method of averaging correlation coefficients used in this study has been to square each coefficient, average these squares, and extract the square root of the average thus obtained. See Garrett, Henry E., op. cit., p. 284.

61. Laird, Donald A., op. cit., p. 212.

sense	180 points
5. Physical and nervous energy	144 points
6. Skill and knowledge	144 points

Defining the Essential Qualities in Terms of Observable Behavior.

The first question was, "What are the tangible actions in a forester's regular work which are dependent upon intelligence for their successful operation?" Each action should refer to but one type of activity. It must be readily recognized, stated in understandable terms, and be particularly characteristic of and applicable to foresters. Care must be taken to see that the actions are significant in relation to the quality described, that it really does measure the quality or trait in question. Also the written action must be of such a nature that it means the same thing to different people all of the time.

For example, the statement, "He readily 'sees the point' in the explanation of a new process," aims to meet these conditions for intelligence. It isn't something vague and mysterious, as is the quality of intelligence. It is readily understood and definite. It is something that a supervisor has many opportunities to observe. Also, it is a significant symptom of intelligence; because if one readily sees the point when a new process is explained it indicates the presence of intelligence.

Other examples of actions which illustrate objective behavior that indicate the presence or absence of intelligence are:

1. He can read and follow written instructions.
2. He effectively analyzes and organizes his work.
3. Etc.

Each of the other five essential qualities (personality, character, scientific trend of mind, etc.) was attacked and defined in the same manner. After a number of samples had been made, experimented with, and re-written for each group the fifteen most promising statements under intelligence, personality, character, scientific trend of mind and common sense along with the twelve best under physical and nervous energy and skill and knowledge were listed in six groups under the heading, "Check list of actions for ascertaining performance." The six groups were labeled from Part I through Part VI in the order named.

Two rating scales were constructed. One may be described as a combination of the descriptive and multiple step types, entitled "Service Record," while the other is a form of the man-to-man type, entitled "Man-to-Man Record." The same descriptive actions were used in each type of rating scale in defining and describing the six essential qualities or parts.

However, there is one important difference in this

connection. The names of the six essential qualities and their brief definitions are retained in the Man-to-Man Record as headings for each group of observable actions. In the Service Record the names and brief definitions were eliminated. These were eliminated in the interest of greater objectivity. It was thought that raters might consciously or unconsciously rate each man on the fifteen actions or items under intelligence in relation to their own definition of intelligence and not on the observable actions given. Different raters might have slightly different opinions as to the meaning of intelligence and these definitions would not necessarily remain static. This would be equally true for the other five parts. It would mean that different raters would not be judging or rating the same things. Neither would a single rater necessarily be rating the same things on two different occasions.

Developing a Method of Recording Comparative Degrees of Behavior.

The difficulty was encountered of finding an accurate method of recording the relative amount of any quality demonstrated by a man being rated, even though different raters agree upon what the action was that was being rated. A method of recording estimates was needed that would mean the same thing to different raters and the same thing to each rater at different times. Because of this difficulty,

it was decided to try two methods, the adjective type and the man-to-man type. These will be developed separately.

It was decided to use a five step range in the adjective scale in order to get a sufficient range of opinion and still not require a too fine discrimination in judgments. Care was taken to avoid such words as good, fair, and poor because of possible bias due to past experience and because of the lack of agreement as to the real meaning of any such word. Highest, high, medium, low, and lowest were the first five words chosen. The rater was instructed to judge each man on the observable actions in relation to the range of performance suggested by these terms. They were defined as, "highest means ideal, the best that could be expected in this position; lowest is the opposite, the least necessary to get by in the position; medium is halfway between or midway; and high and low are intermediate between highest and medium and medium and low, respectively."

These terms were experimentally tried out with a group of ten raters, but the terms did not prove to be entirely satisfactory. Highest and the other terms didn't mean the same to different raters.

The terms finally selected as most satisfactory and included in the final scale consist of "almost always," "usually," "about half of the time," "occasionally," and

"hardly ever." They are used as answers to the action statements, which are in effect questions. For instance, the rater is instructed to use one of these five terms in answer to the question, "His size and appearance attract favorable attention?" Raters have found this method of describing the different degrees of behavior more satisfactory. It is definite, needs very little explanation, and there is much more agreement between raters as to the meaning of the terms.

A man-to-man scheme of rating was tried because it was thought that it might make the judgments more objective because the raters would be forced to judge the ratees on the action statement in comparison with known men instead of with intangible words. This method requires that a master scale of three men be devised for each of the six essential qualities. The three men on the master scale represent the range of performance from the best employee that the rater has known in that position to the poorest employee in that position that the rater has known. The third man is half way between these extremes. These three men listed for each essential quality represent the total range for each quality and are represented on the rating scale by the words maximum, medium and minimum. The raters are then told to keep the master scale of men in mind while rating a man on the actions defining each essential quality.

They are to compare the rates on each action under each essential quality with the three men on the master scale. If the ratee is most like the man pictured as minimum, that should be checked, etc., through the entire rating scale by groups.

Directions to the Rater.

The mechanics of the rating scheme are carefully explained in the directions. Raters are cautioned to consider each item separately. The point that each employee should be judged in relation to the job which he is assigned to do and not in comparison with someone in another position is emphasized.

Raters are cautioned not to guess when checking on any of the items. Directions on one of the scales say, "If you are not reasonably sure from your own experience that the man demonstrates the trait in that amount, leave out that item. No definite number of items need be checked." Raters are directed to report on only those with respect to which they have definite knowledge in order to eliminate guessing, obtain a more careful and objective estimate on the items checked, and to eliminate compromise ratings which are generally rated as average because of inadequate information. Adjustments are made for each of the six groups of items and the unchecked items are given an average value as shown by the checked items in the group.

It should be noted that if the ratings on the items checked are accurate the giving of an average rating on the unchecked items is also accurate. In practice it should surely be as accurate and refined as rating scale conditions warrant. The final score is only intended to place individuals accurately into one of five groups.

### Scoring Methods.

Each item is given points according to the following plan in each of the two rating scales:

Service Record Rating Scale: Man-to-Man Record Rating Scale:

Almost always	12 points	Maximum	12 points
Usually	9 points	Medium	6 points
About half of the time	6 points	Minimum	0 points
Occasionally	3 points		
Hardly ever	0 points		

Scores are totaled for each part or group of items or actions defining major qualities. There is a place for recording the total for each part on the scale. Total possible scores by parts for both types of rating scales may be found under the heading, "Weighing the essential qualities," on page 115.

The total numerical score is found by adding the scores of the six parts. The total possible score is 1008 points.



Scores are totaled for each group of items and if any items have been omitted they are given a value equal to the average value of the checked items. If any items are left out the score for the part can be easily and quickly computed with the following formula, when "X" is the total score for the part.

$$X = \frac{(\text{points on checked items}) \cdot (\text{total no. of items in part})}{(\text{number of items checked})}$$

This formula is developed from the proportion: The total score for the part (X) is to the score of the items checked as the number of items in the part is to the number of items checked.

#### Converting Total Numerical Scores Into Final Ratings of Five Classes.

Total numerical scores are changed into one of five final classes. The limiting points between these classes, or the termini of these classes should be based upon as many total numerical scores as are available.

The ratings upon which computations are based should be made under normal conditions with the group for which the rating scale was intended. Five Forest Supervisors and Assistant Forest Supervisors used the Service Record rating scale in rating the professional personnel. Altogether forty-two ratings were made by them. The distribution of their total scores is shown in Table II.

TABLE II

Distribution of Numerical Scores Made by Forest Supervisors  
and Assistant Forest Supervisors  
in Rating Their Professional Personnel

Class-Intervals	Frequencies
920 - 960	1
880 - 919	1
840 - 879	0
800 - 839	3
760 - 799	3
720 - 759	6
680 - 719	6
640 - 679	6
600 - 639	2
560 - 599	2
520 - 559	2
480 - 519	2
440 - 479	0
400 - 439	2
360 - 399	1
320 - 359	1
	<u>1</u>
	N - 42

Mean -  $670.00 \pm 13.53$

Standard Deviation  $130.00 \pm 9.57$

Although the distribution tends to follow a normal distribution there are too few samples to divide the group accurately on a straight numerical basis. Therefore, the five classes are divided statistically. Briefly the method followed is:

Decide what per cent of the employees are desired in each group. A five class grouping recommended by some<sup>62</sup> is ten per cent in the "A" group, twenty per cent in the "B" group, forty per cent in the "C" group, twenty per cent in the "D" group, and ten per cent in the "E" group. Of course, other letters or names may be used for these groups such as excellent, good, fair, poor, and unsatisfactory.

Class divisions were determined using the above suggested per cent divisions and also according to a five class grouping based entirely on the normal curve.

The statistical method<sup>63</sup> used in determining the groups is briefly as follows:

1. Find the mean and standard deviation of the frequency distribution.
2. Find the length of the standard deviation in each of the five classes. The method is slightly different for the assumed and normal distribution classes. The total distribution is assumed to be six standard deviations.

---

62. Patterson, Donald G., op. cit., p. 365.

63. Clinton, R. J., op. cit.

a. When using the assumed per cent classes, ten, twenty, forty, twenty, and ten per cents, consult a table<sup>64</sup> which gives normal distribution in terms of standard deviations. It shows that the distribution in terms of standard deviations for the five classes, as:

A class

+1.28 standard deviations to +3.00 standard deviations

B class

+ .50 standard deviations to +1.28 standard deviations

C class

- .50 standard deviations to + .50 standard deviations

D class

-1.28 standard deviations to - .50 standard deviations

E class

-3.00 standard deviations to -1.28 standard deviations

b. In a normal distribution the number of the standard deviations for each of the five classes is found by dividing the six standard deviations in the distribution by the five classes, which gives 1.2 standard deviations in each class. By consulting a table<sup>65</sup> which gives the distribution of normal cases in terms of standard deviations, the percentages in each class are found. Results are:

A class 3.5 per cent

C class 45.0 per cent

B class 24.0 per cent

D class 24.0 per cent

E class 3.5 per cent

---

64. Garrett, Henry E., op. cit., p. 110.

65. Ibid.

3. Now that the number of the standard deviations are known for each of the five classes they are applied to the frequency distribution at hand. The division limits for each of the two types of groupings are shown in Table III.

TABLE III

Conversion Table for Changing Total Numerical Scores  
to Class Scores

Classes	Class limits based on 10, 20, 40, 20, and 10 per cent distribution of classes	Class limits based on a normal distribution
A	836 and up	904 and up
B	735 to 836	748 to 904
C	605 to 735	592 to 748
D	504 to 605	436 to 592
E	280 to 504	280 to 436

Providing an Opportunity for Special Remarks.

In order to cover any points missed or insufficiently stressed in the rating scale two spaces for additional remarks have been provided. The first asks, "What outstanding qualifications or deficiencies, if any, does he have that would qualify him or disbar him in another or a higher position?"

The second space makes no suggestions as to the type of information that should be included, but leaves it entirely to the discretion of the rater.

Final Rating Scales.

All of the parts were assembled after many changes and alterations and the two finished rating scales follow:

SERVICE RECORD

Employee \_\_\_\_\_  
 Position \_\_\_\_\_  
 Score \_\_\_\_\_

Rater \_\_\_\_\_  
 Position \_\_\_\_\_  
 Date \_\_\_\_\_

DIRECTIONS. Following is a list of actions which are to be used in describing a man. Place a check (✓) after the statement in the column that properly describes or fits the man that is being checked, on the basis of his performance on the job to which he is assigned. The meaning of the terms used are self evident. Of course, no one man demonstrates flawlessly all of these actions listed, but some approach this ideal to a much greater extent than do others. However, carefully consider each action in relation to the wide range of performance suggested. Do not guess; if you are not reasonably sure from your own experience that the man demonstrates the action or trait in that amount in his assigned job, omit that item. No definite number of items need be checked.

Has demonstrated  
 performance ability  
 to the following extent:

Check List of Actions for  
 Ascertaining Performance:

Nearly always	Usually	About half of the time	Occasionally	Hardly ever
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____

Part I.

1. He readily "sees the point" in an explanation of a new process.
2. He ably meets complex and unexpected individual and group social situations.
3. He is able to break a situation down into its parts and reassemble it, as illustrated by his being able to take an unfamiliar piece of machinery apart, see how it works and put it together again.

	Nearly always	Usually	About half of the time	Occasionally	Hardly ever
4. He changes methods and makes quick adjustments in order to meet practical problems.	—	—	—	—	—
5. He can read and follow written instructions.	—	—	—	—	—
6. He effectively analyzes and organizes his work.	—	—	—	—	—
7. He is resourceful in meeting unexpected situations.	—	—	—	—	—
8. He is able to see relationships between different situations.	—	—	—	—	—
9. He is able to recognize the special abilities and limitations of others.	—	—	—	—	—
10. He is able to pick out the most important problems of his work.	—	—	—	—	—
11. He is actively interested in solving the problems of his work.	—	—	—	—	—
12. He gives prompt and definite decisions.	—	—	—	—	—
13. He is able to fore-see and forestall dissension among the members of the group.	—	—	—	—	—
14. He is able to fore-see future needs and acts accordingly. For example, in the opening up of a spring he also prepares a pump site that may be used in future fire control.	—	—	—	—	—
15. He carefully studies his limitations and weaknesses and proceeds systematically to remove them.	—	—	—	—	—

Part I score \_\_\_\_\_



(Remember, no definite number of items need be checked.)

### Part II.

1. His size and appearance attract favorable attention.
2. His voice and diction command attention and respect.
3. His clothing is always appropriate for the occasion.
4. He is neat and clean in body and habits.
5. He uses good judgment and tact in social, personal, and professional activities.
6. He has no unpleasant mannerisms.
7. He has a keen sense of humor and is not readily offended.
8. He displays poise in meeting difficult situations.
9. He takes a positive and aggressive attitude toward personal and professional problems.
10. He shows that he knows what he is going to do and how he is going to do it.
11. He inspires loyalty among his fellows.
12. He appears cheerful and cordial.
13. He at all times displays a calm but firm feeling of dignity.
14. He gets enthusiastic and optimistic responses.
15. He appears interested in people and is effective in conversation.

Nearly always	Usually	About half of the time	Occasionally	Hardly ever
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—

Part II score \_\_\_\_\_

### Part III.

1. He is genuinely interested in people and seeks to know them better.

\_\_\_\_\_

	Nearly always	Usually	About half of the time	Occasionally	Hardly ever
2. He is never snobbish or patronizing.	—	—	—	—	—
3. He gives due credit for suggestions and takes credit only for work actually done.	—	—	—	—	—
4. He acknowledges his mistakes and assumes the blame for his own actions.	—	—	—	—	—
5. He conceives his position as an opportunity for unlimited constructive endeavor and service rather than a display of personal power.	—	—	—	—	—
6. He considers it an obligation to his profession and society to make known new theories and findings rather than a means to personal power.	—	—	—	—	—
7. He gives loyalty to high ideals; honor, justice, truth, and reliability to his group and friends.	—	—	—	—	—
8. He is loyal to co-workers; he keeps silent rather than criticize them adversely to outsiders.	—	—	—	—	—
9. He leads his men instead of just "bossing" them.	—	—	—	—	—
10. He assumes responsibility to his employer and to his men.	—	—	—	—	—
11. He sets an example as a sincere and desirable force for good.	—	—	—	—	—
12. He can be depended upon to do his work properly and conscientiously.	—	—	—	—	—
13. He is a square shooter. He doesn't bluff.	—	—	—	—	—
14. He is just and sympathetic in his relations with subordinates and others.	—	—	—	—	—
15. He inspires confidence through his self-control and impartial attitude.	—	—	—	—	—
Part III score	—	—	—	—	—



	Nearly always	Usually	About half of the time	Occasionally	Hardly ever
14. He adopts wise policies which insure the greatest good for the greatest number over an <u>expected period of time.</u>	_____	_____	_____	_____	_____
15. He exhibits a high degree of self-direction and maintains a critical attitude towards his own work, <u>methods, and results achieved.</u>	_____	_____	_____	_____	_____

Part IV score \_\_\_\_\_

Part V.

1. He exhibits good physical health.	_____	_____	_____	_____	_____
2. He keeps troubles and worries from interfering with his work.	_____	_____	_____	_____	_____
3. He has the ability to put forth concentrated effort for comparatively long periods of time.	_____	_____	_____	_____	_____
4. He has the habit of hard and effective work.	_____	_____	_____	_____	_____
5. He works rapidly and sticks to jobs until they are finished.	_____	_____	_____	_____	_____
6. He is direct and definite in his general bearing.	_____	_____	_____	_____	_____
7. He is prompt in submitting all written materials such as reports.	_____	_____	_____	_____	_____
8. He is a self-starter.	_____	_____	_____	_____	_____
9. He starts things, plans activities, and starts others.	_____	_____	_____	_____	_____
10. He is able to see what needs to be done without being told what to do.	_____	_____	_____	_____	_____
11. He checks to see that his orders are carried out to insure that they will be carried out completely.	_____	_____	_____	_____	_____
12. He has force to carry out a program.	_____	_____	_____	_____	_____

Part V score \_\_\_\_\_

Nearly always	Usually	About half of the time	Occasionally	Hardly ever
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—
—	—	—	—	—

Part VI.

1. He keeps well informed on current professional literature.
2. He visits other projects and demonstrations to see other ways of doing things.
3. He meets with conferences of other leaders to gain advantage of exchange of ideas.
4. He is a member of professional organizations.
5. He seeks carefully to become really expert in various kinds of work.
6. He has achieved technical mastery of the work to be done; therefore he can explain "why" things are done and thereby help build loyalty and confidence.
7. He possesses enough grasp of all situations under his control to give wise guidance to directive effort.
8. He is able to recognize good work because he is an expert in doing the work himself.
9. He has had experience in varied lines of work and therefore knows how much to expect of his men.
10. He is able to break in and train inexperienced men.
11. He finishes a job according to the standards required.
12. His work is accurate and neat.

Part VI score \_\_\_\_\_

MAN-TO-MAN RECORD

Employee \_\_\_\_\_  
 Position \_\_\_\_\_  
 Score \_\_\_\_\_

Rater \_\_\_\_\_  
 Position \_\_\_\_\_  
 Date \_\_\_\_\_

**DIRECTIONS.** You are going to judge or rate a man's ability and success on the job. It is assumed that all of the qualities needed for successful foresters are included in the following list: Intelligence, personality, character, scientific trend of mind and common sense, physical and nervous energy, and skill and knowledge. The meaning of these terms are given on the following page. In the spaces provided check (✓) the employee so that his action for each point is accurately evaluated.

Before rating or evaluating the individual, for any particular job classification, visualize all of the employees that you have known while they were doing work similar to that required in the particular position in question. For example, when rating a district ranger consider all of the district rangers with whom you have had experience. Consider, first, the quality of intelligence only; do not include personality, character, etc.

Recall the man who excelled all of the others in intelligence. Now picture the poorest one that you have known, the one who conspicuously lacked intellectual capacity. Next, recall a man who ranks about midway between the two extremes, or is average. Write down the names of these three men in the spaces provided. (The list of names should be destroyed after the rating.) These three men listed represent the range in intellectual capacity represented by the words maximum, medium, and minimum.

Keep these three men clearly in mind and their names within sight while rating a man on the points under intelligence. Compare the man to be rated on each point under intelligence to the three men listed. If he is most like the man pictured as minimum, check that point.

After completing the rating on intellectual capacity start all over again. Now consider the quality of personality only. Again recall three men, in a similar position, who represent the widest range in the quality of personality. These three men need not be the same ones considered under intelligence. List at the top the man who makes the best impression and at the bottom the one who habitually makes the least satisfactory impression. Now classify the man who ranks intermediate, as you did when considering

intelligence.

As before, check the man to be rated against the three men used as a standard and rate accordingly.

Before taking up each of the other four characteristics, such as skill and knowledge, conscientiously write down the names of the three men with whom you will compare the man to be rated.

<u>INTELLIGENCE</u> (Intellectual capacity which determines how well we do things.)	Maximum _____
	Medium _____
	Minimum _____
<u>PERSONALITY</u> (The way an individual impresses other folks.)	Maximum _____
	Medium _____
	Minimum _____
<u>CHARACTER</u> (What a man is on the inside.)	Maximum _____
	Medium _____
	Minimum _____
<u>SCIENTIFIC TREND OF MIND AND COMMON SENSE</u> (Solving problems impartially and wisely in the light of all available evidence.)	Maximum _____
	Medium _____
	Minimum _____
<u>PHYSICAL AND NERVOUS ENERGY</u> (Drive, endurance, and vigor of body and mind.)	Maximum _____
	Medium _____
	Minimum _____
<u>SKILL AND KNOWLEDGE</u> (Ability to do things resulting from training, education, and experience.)	Maximum _____
	Medium _____
	Minimum _____

(Destroy after using)

(Be sure and read the instructions carefully and fill out the attached sheet before beginning on this page. Any questions that you can't honestly answer from your own experience please leave unmarked.)

	Max- imum	Med- ium	Min- imum
1. <u>INTELLIGENCE</u> (Intellectual capacity which determines how well we do things.)			
1. He readily "sees the point" in an explanation of a new process.	_____	_____	_____
2. He ably meets complex and unexpected individual and group social situations.	_____	_____	_____
3. He is able to break a situation down into its parts and reassemble it, as illustrated by his being able to take an unfamiliar piece of machinery apart, see how it works and put it together again.	_____	_____	_____
4. He changes methods and makes quick adjustments in order to meet practical problems.	_____	_____	_____
5. He can read and follow written instructions.	_____	_____	_____
6. He effectively analyzes and organizes his work.	_____	_____	_____
7. He is resourceful in meeting unexpected situations.	_____	_____	_____
8. He is able to see relationships between different situations.	_____	_____	_____
9. He is able to recognize the special abilities and limitations of others.	_____	_____	_____
10. He is able to pick out the most important problems in his work.	_____	_____	_____
11. He is actively interested in solving the problems of his work.	_____	_____	_____
12. He gives prompt and definite decisions.	_____	_____	_____
13. He is able to fore-see and fore-stall dissension among the members of the group.	_____	_____	_____
14. He is able to fore-see future needs and act accordingly. For example, in the opening up of a spring he also prepares a pump site that may be used in future fire control.	_____	_____	_____
15. He carefully studies his limitations and weaknesses and proceeds systematically to remove them.	_____	_____	_____

Part I score \_\_\_\_\_



	Max- imum	Med- ium	Min- imum
II. <u>PERSONALITY</u> (The way an individual impresses other folks.)			
1. His size and appearance attract favorable attention.			
2. His voice and diction command attention and respect.			
3. His clothing is always appropriate for the occasion.			
4. He is neat and clean in body and habits.			
5. He uses good judgment and tact in social, personal, and professional activities.			
6. He has no unpleasant mannerisms.			
7. He has a keen sense of humor and is not readily offended.			
8. He displays poise in meeting difficult situations.			
9. He takes a positive and aggressive attitude toward personal and professional problems.			
10. He shows that he knows what he is going to do and how he is going to do it.			
11. He inspires loyalty among his fellows.			
12. He appears cheerful and cordial.			
13. He at all times displays a calm but firm feeling of dignity.			
14. He gets enthusiastic and optimistic responses.			
15. He appears interested in people and is effective in conversation.			

Part II score \_\_\_\_\_

III. CHARACTER (What a man is on the inside.)

1. He is genuinely interested in people and seeks to know them better.			
2. He is never snobbish or patronizing.			
3. He gives due credit for suggestions and takes credit only for work actually done.			
4. He acknowledges his mistakes and assumes the blame for his own actions.			
5. He conceives his position as an opportunity for unlimited constructive endeavor and service rather than a display of personal power.			

	Max- imum	Med- ium	Min- imum
6. He considers it an obligation to his profession and society to make known new theories and findings rather than a means to personal power. - - - - -	_____	_____	_____
7. He gives loyalty to high ideals: honor, justice, truth, and reliability to his group and friends. - - - - -	_____	_____	_____
8. He is loyal to co-workers; he keeps silent rather than criticize them adversely to outsiders. - - - - -	_____	_____	_____
9. He leads his men instead of just "bossing" them. - - - - -	_____	_____	_____
10. He assumes responsibility to his employer and to his men. - - - - -	_____	_____	_____
11. He sets an example as a sincere and desirable force for good. - - - - -	_____	_____	_____
12. He can be depended upon to do his work properly and conscientiously. - - - - -	_____	_____	_____
13. He is a square shooter. He doesn't bluff. - - - - -	_____	_____	_____
14. He is just and sympathetic in his relations with subordinates and others. - - - - -	_____	_____	_____
15. He inspires confidence through his self-control and impartial attitude. - - - - -	_____	_____	_____

Part III score \_\_\_\_\_

- IV. SCIENTIFIC TREND OF MIND AND COMMON SENSE (Solving problems impartially and wisely in the light of all available evidence.)
1. He readily senses the presence of a perplexing problem and its nature. - - - - -
  2. He examines critically all proposed solutions of problems until all of the facts are gathered. - - - - -
  3. He casts aside any plans found invalid. - - - - -
  4. He suspends judgment on problems until all of the facts are gathered. - - - - -
  5. He re-checks conclusions to test their validity. - - - - -
  6. He is able to keep a problem in mind when hunting for a solution. - - - - -
  7. He is open-minded and is able to use the ideas of others. - - - - -
  8. He gives decisions on the basis of fair objective reasons rather than personal bias or interest. - - - - -

	Max- imum	Med- ium	Min- imum
9. He reprimands properly because he does so only after a systematic study of the problem.	_____	_____	_____
10. He develops team-work by placing men in the right positions, and allocating responsibility for results by previous planning.	_____	_____	_____
11. He presents both sides of a question and encourages decisions on the relative worth of the evidence.	_____	_____	_____
12. His knowledge of cause and effect relationships make him confident of the outcome of his endeavors.	_____	_____	_____
13. His fore-sight is accurate and dependable because it is founded upon accurate observations and information.	_____	_____	_____
14. He adopts wise policies which insure the greatest good for the greatest number over an expected period of time.	_____	_____	_____
15. He exhibits a high degree of self-direction and maintains a critical attitude towards his own work, methods, and results achieved.	_____	_____	_____

Part IV score \_\_\_\_\_

V. <u>PHYSICAL AND NERVOUS ENERGY</u> (Drive, endurance, and vigor of body and mind.)			
1. He exhibits good physical health.	_____	_____	_____
2. He keeps troubles and worries from interfering with his work.	_____	_____	_____
3. He has the ability to put forth concentrated effort for comparatively long periods of time.	_____	_____	_____
4. He has the habit of hard and effective work.	_____	_____	_____
5. He works rapidly and sticks to jobs until they are finished.	_____	_____	_____
6. He is direct and definite in his general bearing.	_____	_____	_____
7. He is prompt in submitting all written materials such as reports.	_____	_____	_____
8. He is a self-starter.	_____	_____	_____
9. He starts things, plans activities, and starts others.	_____	_____	_____
10. He is able to see what needs to be done without being told what to do.	_____	_____	_____

	Max- imum	Med- ium	Min- imum
--	--------------	-------------	--------------

- |                                                                                                          |  |  |  |
|----------------------------------------------------------------------------------------------------------|--|--|--|
| 11. He checks to see that his orders are carried out to insure that they will be carried out completely. |  |  |  |
| 12. He has force to carry out a program.                                                                 |  |  |  |

Part V score \_\_\_\_\_

# VI. SKILL AND KNOWLEDGE (Ability to do things resulting from training, education, and experience.)

- |                                                                                                                                                             |  |  |  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|--|--|--|
| 1. He keeps well informed on current professional literature.                                                                                               |  |  |  |
| 2. He visits other projects and demonstrations to see other ways of doing things.                                                                           |  |  |  |
| 3. He meets with conferences of other leaders to gain advantage of exchange of ideas.                                                                       |  |  |  |
| 4. He is a member of professional organizations.                                                                                                            |  |  |  |
| 5. He seeks carefully to become really expert in various lines of work.                                                                                     |  |  |  |
| 6. He has achieved technical mastery of the work to be done; therefore, he can explain "why" things are done and thereby help build loyalty and confidence. |  |  |  |
| 7. He possesses enough grasp of all situations under his control to give wise guidance to directive effort.                                                 |  |  |  |
| 8. He is able to recognize good work because he is an expert in doing the work himself.                                                                     |  |  |  |
| 9. He has had experience in varied lines of work and therefore knows how much to expect of his men.                                                         |  |  |  |
| 10. He is able to break in and train inexperienced men.                                                                                                     |  |  |  |
| 11. He finishes a job according to the standards required.                                                                                                  |  |  |  |
| 12. His work is accurate and neat.                                                                                                                          |  |  |  |

Part VI score \_\_\_\_\_

## CHAPTER XIII

### TESTING THE RATING SCALE

## CHAPTER XIII

### TESTING THE RATING SCALE

Now that two rating scales have been constructed, several questions arise. Will they measure foresters efficiently? How do they compare with other rating scales? How can they be improved?

An attempt is made to answer these questions through an analysis of three experiments. The first two experiments were conducted in the School of Forestry at Oregon State College while the third was carried on in the United States Forest Service. It is acknowledged that results obtained in experiments conducted other than in the Forest Service cannot be conclusively relied upon to forecast results in the Forest Service. However, it is assumed that forestry school professors, seniors, and graduate students are similar to foresters in the Forest Service and that results in certain types of tests would be indicative of what could be expected under actual Forest Service conditions.

#### Comparing the Service Record and Man-to-Man Record Rating Scales.

An experiment was conducted to determine which is more satisfactory, the Service Record Rating Scale or the Man-to-Man Record Rating Scale. Nine upper division and graduate students and a professor in the School of Forestry rated

a man known to the entire group. Before attempting to do the rating, two one-hour meetings were used in discussing and familiarizing the group with rating scales. The final scores in this experiment appear below in Table IV.

TABLE IV

Ratings Made by Nine Students and One Professor  
in the School of Forestry of One Man with the  
Service Record and Man-to-Man Record Rating Scales

Rater	Service Record Score	Man-to-Man Record Score
SG 1	780	718
SG 2	788	586
SG 3	592	506
SG 4	783	760
SG 5	714	622
SG 6	744	608
SG 7	794	752
SG 8	633	616
SG 9	854	866
SG 10	702	696
Mean	738.4	673
Standard Deviation	75.50	99.58

A standard deviation of 75.50 on the Service Record and a standard deviation of 99.58 on the Man-to-Man Record indicates that there is approximately one third less variation between raters on the Service Record rating scale than the Man-to-Man Record rating scale when they are used on the same man.

Perhaps the main reason for the greater variation

between raters on the Man-to-Man Record rating scale was because of the difficulty encountered in making the master scale. The raters were not able to satisfactorily list three men representing maximum, medium, and minimum efficiency for each of the six main divisions. This may have been due to inexperience and limited acquaintance of most of the raters.

From this experiment it was concluded that the Service Record rating scale is the more accurate measuring instrument of the two and no further experiments were conducted with the Man-to-Man rating scale.

#### Rating Scale Testing in the School of Forestry.

This experiment attempts to compare three rating scales: Form 3200,<sup>1</sup> which is used in the Forest Service, the Probst Service Report, and the Service Record Rating Scale.

Four professors in the School of Forestry were asked to rate ten men, whom they had had in classes. These men all graduated in June, 1938. The professors rated these men with each of the above named rating scales in January and again in March, 1939. The resultant scores of this

---

1. The rating scale used in the Forest Service is entitled Service Rating Form of which there are two similar forms. Form 3200 is used in this experiment. It will be referred to by number to avoid confusion that might result due to the similarity of titles of the three rating scales named.



experiment are shown in Tables V and VI, in the appendix.

Assuming that the ratees are of about the same general value in one month as they are in another month, each professor's ratings of these same ratees should be consistent from January until March. With this assumption the reliability or consistency with which each of the ratings scales is used at different times is determined. This is found by determining the amount of agreement between the individual ratings given the same ratees in January and those given in March by each of the professors. These agreements are evaluated in terms of the coefficient of correlation. Yoder states that a correlation of plus .75 is needed.<sup>2</sup> These correlations follow:

	Form 3200	Probst Service Report	Service Record
Professor # 1	1.00	.89	.78
Professor # 2	.87	.92	.79
Professor # 3	.81	.80	.84
Professor # 4	.91	.94	.82
Average <sup>3</sup>	.90	.89	.81

The average coefficients, .90 and .89, for the Form 3200 and Probst Service Report ratings scales, respectively, indicate that they were satisfactorily consistent or reliable in this trial. The average coefficient of .81 for

---

2. Yoder, Dale, op. cit., p. 270.

3. Garrett, Henry E., op. cit., p. 284.

the Service Record rating scale indicates that it was about fifteen per cent less consistent than the other two rating scales.

The lower reliability of the Service Record rating scale in this experiment may be attributed to the fact that on the whole only about one half of the items in this rating scale were answered by the professors. The reason advanced for this was that the Service Record scale calls for special information about the ratees which professors are not in a position to know, but which applies in the Forest Service. It was remarked by one rater that he would know this information about his subordinates if he were in the Forest Service. It is possible that this rating scale would prove to be more reliable if the raters were better acquainted with the ratees.

It is concluded that the Service Record rating scale is not sufficiently reliable while Form 3200 and the Probst Service Record rating scales are sufficiently reliable when forestry students are rated by professors. However, it is felt that the lack of sufficient information about the ratees abnormally affected the Service Record results and the true consistency of the scale as shown by a rater's ratings at different times in the Forest Service can only be determined by actual experiments in the Forest Service.

### Testing the Rating Scales in the Forest Service.

As previously mentioned, an attempt has been made by employees of the Federal Government to make Form 3200, the rating scale used in the U. S. Forest Service, more objective by supplementing it with a graphic rating scale entitled, "Guide For Using Service Rating Form 3200."<sup>4</sup> In this experiment the "Guide For Using Form 3200" Rating Scale, a copy of which is in the appendix, is compared with the Probst Service Report and the Service Record Rating Scales.

It was planned to conduct the experiment under normal conditions. Therefore, all instructions and communications were handled through regular administrative channels.

Instructions were sent to three National Forests asking the Forest Supervisor and Assistant Forest Supervisor to rate their subordinates who were professional foresters with each of the three above mentioned rating scales.

Altogether twenty-five men were rated. They were each given a number ranging from one to twenty-five, inclusive, and they, henceforth, will be referred to by number. The raters will be identified according to the following plan:

Forest Supervisors ----- B, C, and E.

Assistant Forest Supervisors A, D, and F.

Executive Assistant ----- G.

---

4. For convenience this rating scale will be referred to as "Guide for Using Form 3200."

The experiment was carried out as planned on the first two Forests. Raters A and B rated Ratees one to ten and Raters C and D rated Ratees eleven to seventeen on all three scales. On the third, Forest Rater F completed the rating on Ratees eighteen to twenty-five as planned but Rater E was transferred after rating Ratees eighteen to twenty-five on the "Guide For Using Form 3200" Rating Scale and before he had rated the men with either of the other two rating scales. An attempt was made to have Rater G take the place of Rater E but he was unable to rate the men with any of the rating scales except the Probst Service Record because he was not familiar with their field work.

This means that Raters A, B, C, D, and F rated their respective subordinates with all three rating scales making forty-two ratings on the twenty-five men. The most of the statistical analysis will be based upon these ratings and any time ratings by either Rater E or G are used it will be specifically mentioned.

Rating scores made by these seven raters on the twenty-five ratees are shown in Table VII in the appendix. These ratings will now be analyzed.

1. Distribution of ratings. When several men are rated, their total scores should be scattered, as opposed to being bunched together, in order to differentiate between the men and to segregate them into the several

classes of ability as shown by performance. In order to show the distribution of the ratings the total numerical ratings were converted into the final class ratings as provided in each rating scale.<sup>5</sup> The results are compared and for convenience the class ratings of excellent, very good, good, fair, and unsatisfactory obtained with the "Guide For Using Form 3200" rating scale are shown under the letters A, B, C, D, and E as used in the other two rating scales.<sup>6</sup> These final class ratings together with the total spread on each rating scale as shown in the forty-two ratings on the twenty-five men follow:

	Guide for <sup>7</sup> Form 3200	Probst Service Report	Service Record
Total spread	81 to 93.5	-52 to 76	331 to 959
Number of ratings	42	42	42
Number of A ratings	6	3	2
Number of B ratings	32	6	10
Number of C ratings	4	22	20
Number of D ratings	0	9	4
Number of E ratings	0	2	5

5. Conversion tables for changing total numerical scores into final ratings are shown in Table XI of the appendix. The table for converting Form 3200 scores into numerical ratings with the highest scores instead of the low scores as shown on Form 3200 representing the highest class is not shown.
6. These class Ratings are shown in Table VIII in the appendix.
7. Guide For Form 3200 refers to Guide For Using Service Rating Scale Form 3200.

Normally the greatest number of ratings should be concentrated in the middle or "C" class with progressively fewer in each step down to the "A" and "E" classes.<sup>8</sup> Both the Probst Service Report and the Service Record results show a concentration of ratings in the middle class with fewer in the end classes. However, the "Guide For Using Form 3200" shows the preponderance of ratings in the B instead of the C class and no ratings in the lower two classes.

Figure 3 graphically shows this distribution for the three rating scales and the failure of the "Guide For Using Form 3200" to segregate men into classes because of a lack of spread in the ratings.

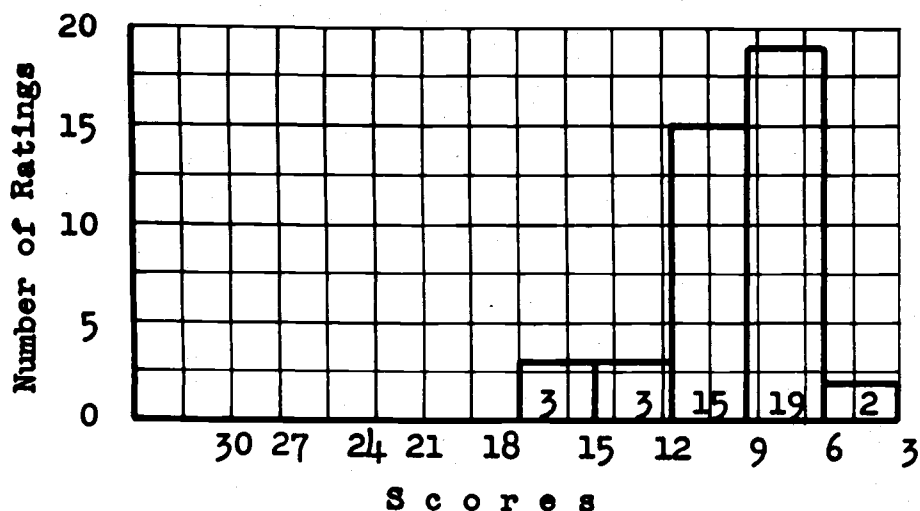
Also theoretically, in a five class rating system such as these being tested, the middle or C class should be the normal mean or average of the ratings. By translating the letter ratings into numerical equivalents according to this plan--

Letter ratings	A	B	C	D	E
Numerical equivalents	1	2	9	6	3

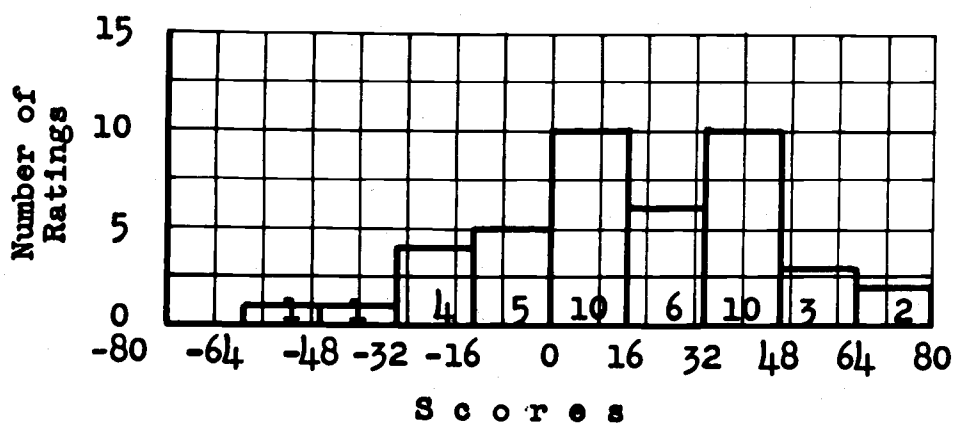
then the theoretical mean should be exactly six. Of course, small groups are not expected to produce normal

---

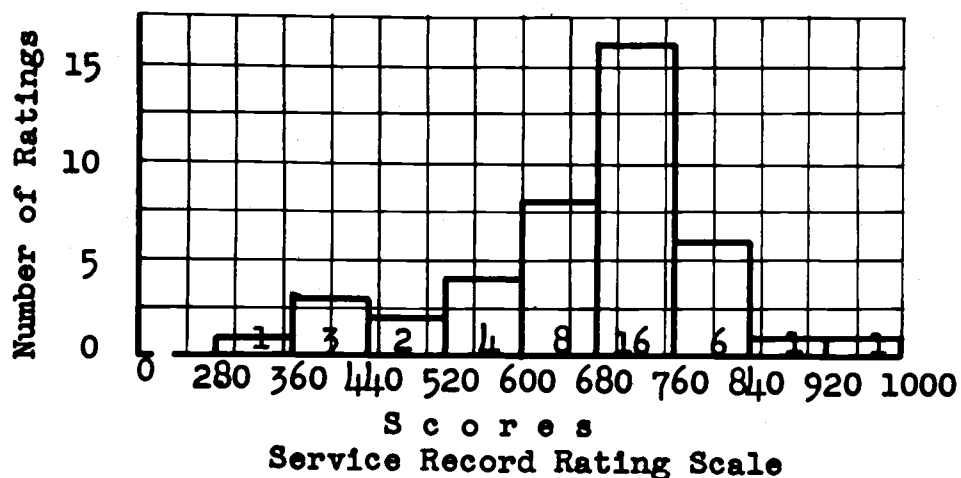
8. On a large grouping Probst says that the normal distribution by classes for his rating scale is 4.7 per cent in A class, 17.3 per cent in B class, 55.2 per cent in C class, 17.7 per cent in D class, and 5.1 per cent in E class.



Guide to Form 3200 Rating Scale



Probst Service Report Rating Scale



Service Record Rating Scale

Figure 3. Distribution of 42 ratings on 25 men in the Forest Service made with "Guide to Form 3200," "Probst Service Report," and "Service Record" rating scales. (Numbers in columns indicate number of ratings.)

averages or normal distributions, but the distribution of a small group may indicate the distribution trend of a larger but similar group.

The mean, or average, of the forty-two ratings on the twenty-five men is as follows for the three types of rating scales:

Guide for Using Form 3200	9.14
Probst Service Report	5.98
Service Record	5.98

The above three examples indicate that the Probst Service Report and the Service Record Rating Scales segregate the rates into classes on the basis of a fairly normal distribution. At the same time the Guide For Using Form 3200 Rating Scale concentrates the final scores in the next to the highest class. Only four final scores are in the middle class and normally this class should contain the largest number of final scores.<sup>9</sup>

The Guide For Using Form 3200 Rating Scale fails to segregate the final scores into a distribution such as is normally expected. Two reasons for this are cited. The first reason is the lack of objectivity in the items. They are defined in such broad terms as "General dependability; accuracy" and "Knowledge of duties and related information."

---

9. United States Department of Agriculture, Efficiency Rating Manual, p. 4.



General terms are then used to show the range of the defined qualities. An example is, "Uses excellent procedure, good procedure, fair procedure, and poor procedure used." It is natural for raters to tend to rate too high<sup>10</sup> and when a rater cannot rate specifically, both as to qualities and degree, he tends to give a general and high rating.

The second reason for the lack of proper distribution of the ratings is the limited number of items rated. There are fifteen items divided into three groups. Increasing the number of items and groups would increase the range and the possibility of a wider distribution of final ratings.

It is concluded that the Guide For Using Form 3200 Rating Scale does not divide or segregate the final scores into the classes as normally expected. The Probst Service Record and the Service Record Rating Scales satisfactorily segregate the final scores.

2. Reliability of the ratings. Reliability is the consistency with which the three rating scales gauge the ability of foresters. Two statistical methods were used in checking the reliability.

First the split-half method<sup>11</sup> of comparison was used in which the total score of odd numbered items were compared with the total score of the even numbered items for each

---

10. Ibid., p. 7.

11. Garrett, Henry E., op. cit., p. 318-19.

rater. Different numbers of items were checked by the raters for the different ratees and in order to make them comparable, totals for both the odds and evens were divided by the number of items checked to get the average. Their averages were then compared. This test was not applicable to results obtained with the Probst Service Report rating scale. Total and average scores obtained with the other two rating scales are shown in Tables IX and X of the appendix.

Results found by correlating the score of odd items against the score of the even items for each rater with the "Guide For Using Form 3200" and "Service Record" rating scales follow:

Rater	Men	Correlation by Split-half Method	
		"Guide to Form 3200"	Service Record
A	1 to 10	.96	.99
B	1 to 10	.96	.96
C	11 to 17	.97	.98
D	11 to 17	.91	.94
F	18 to 25	.74	.99
Average <sup>12</sup>		.91	.98

These average correlations of .91 and .98 indicate a high consistency with both rating scales although the averages indicate that the Service Record Rating Scale correlations are eighty per cent better than chance while

12. Ibid., p. 284.

the "Guide For Using Form 3200" correlations are fifty per cent better than chance.<sup>13</sup>

These correlations may be interpreted to mean that the scores on one half of the items consistently arrange the rates in approximately the same rank order as the scores on the half of the items. Although both rating scales show satisfactory consistency, the Service Record Rating Scale is twenty-one per cent more consistent than the Guide For Using Form 3200 Rating Scale.

The other statistical method of checking reliability used was to compare the ratings of different raters on the same men. Close agreement between different raters on the same men is additional evidence of reliability. These results follow:

Agreement between:	Guide For Form 3200	Probst Report	Service Record
Raters A and B on men # 1 to 10	.72	.59	.66
Raters C and D on men # 11 to 17	.61	.88	.89
Raters E and F on men # 18 to 25	.36		
Raters G and F on men # 18 to 25		.85	
Average <sup>14</sup>	.59	.78	.78

Average correlations of .78 between the different raters on the same men using the Probst Service Report and

---

13. Laird, Donald, A., op. cit., p. 212.

14. Garrett, Henry E., op. cit., p. 284.

the Service Record Rating Scales indicate a fairly close agreement. However, a correlation of .59 on the Guide For Using Form 3200 Rating Scale indicates quite low agreement.

On the basis of these two tests the Service Record Rating Scale meets the requirements of consistency or reliability. Additional comparisons need to be made with the Probst Service Report but what evidence there is indicates that it is equally as reliable as the Service Record Rating Scale. The consistency of the Guide For Using Form 3200 Rating Scale is questionable inasmuch as it showed a low correlation in one of the two tests.

3. Comparison of ratings. It has already been shown that final scores are distributed quite normally by the Probst Service Report and the Service Record Rating Scales while the final scores on the Guide For Using Form 3200 Rating Scale tend to cluster around the next to the highest class. Because of these differences the individual ratings were examined and compared.

The final class ratings, which are shown in Table VIII of the appendix, were checked. First, the final scores of the Guide For Using Form 3200 Rating Scale were examined and compared with the scores obtained with the Probst Service Report Rating Scale to see how many ratees received the same final class rating with both rating scales, how many ratees received final class scores which varied one

step or class, two classes, or three classes with each rater. For instance, a variation of one class is when a ratee is put in the highest class with one rating scale and the next to the highest class with the other rating scale by the same rater. Next, the final scores of the Guide For Using Form 3200 Rating Scale were compared to the final scores obtained with the Service Record Rating Scale in the same manner. Then in the same way the final scores obtained with the Probst Service Report and Service Record Rating Scale were compared. In all cases the scores compared were made by the same rater on the same ratee. The results of these comparisons follow:

Agreement between:	<u>No. of ratees given class ratings which were:</u>				
	<u>Same Class</u>	<u>One step differ- ence</u>	<u>Two step differ- ence</u>	<u>Three step differ- ence</u>	<u>No. rat- ings</u>
Guide For Using Form 3200 and Probst Service Report	8	23	11	0	42
Guide For Using Form 3200 and Service Record	9	21	9	3	42
Probst Service Report and Service Record	18	23	1	0	42

It is very easy to have a one step or one class difference with little actual difference in the final scores. For example, one rating may be one point below the dividing

line between two classes while another rating might be one point above the dividing line. This would show a difference of one class when in reality the two final scores would be separated by only two points. Therefore, a difference of one step or class is not considered a serious disagreement.

There is very close agreement between the Probst Service Report and Service Record Rating Scales because only one out of forty-two final scores varies more than one class. The final class scores obtained with Guide For Using Form 3200 Rating Scale varies quite radically from the Probst Service Report and the Service Record Rating Scales because eleven and twelve, respectively, of the forty-two final class ratings vary two classes or more.

Some of the individual ratee reports were examined where the final class score obtained with the Guide For Using of Form 3200 Rating Scale differed from the final class scores obtained with the other two rating scales. The results of three of these samples follow:

Guide For Using Form 3200 -- Very good (next to  
highest class)

Probst Service Report ----- D (next to lowest  
class)

Service Record ----- C (middle class)

While this man was given a final rating one class above the average with the Guide For Using Form 3200, both

raters checked the following items as true. Only part of the unfavorable items checked by both raters are reproduced.

**Probst Service Report:**

Slow moving  
Often needs prodding  
Might often use better judgment  
Needs considerable supervision

**Service Record:**

About one half of the time:

Effectively analyzes and organizes his work.  
Can be depended upon to do his work properly and conscientiously.  
Able to see what needs to be done without being told what to do.  
Has force to carry out a program.

Ratee 17 was given the following final class ratings by Raters C and D:

Guide For Using Form 3200 -- Good (middle class)

Probst Service Report ----- E (lowest class)

Service Record ----- E (lowest class)

Some of the items checked as true by both raters concerning this man follow:

**Probst Service Report:**

Lazy  
Does not do his share of the work  
Must generally be told what to do  
Often does careless work  
Poor technical training for the work

**Service Record:**

Seldom finishes a job according to the standards required.  
Seldom can be depended upon to do his work properly and conscientiously.

Seldom has force to carry out a program.  
 Seldom starts things or plans activities.  
 Seldom possesses good physical health.

Rater B1 was given the following final class ratings  
 by Rater F:

Guide For Using Form 3200 -- Very Good (next to highest class)

Probst Service Report ----- D (next to lowest class)

Service Record ----- E (lowest class)

Some of the items checked as true concerning this man follow:

**Probst Service Report:**

Discipline too harsh  
 Might often use better judgment  
 Easily rattled  
 Learns new work slowly  
 Often does careless work  
 Poor technical training for the work  
 Does not plan or lay out work properly

**Service Record:**

Seldom displays poise in meeting difficult situations.  
 Seldom appears cheerful and cordial.  
 Seldom just and sympathetic in his relations with subordinates and others.  
 Seldom assumes responsibility to his employers and to his men.  
 Seldom able to break in and train inexperienced men.

4. Validity of the ratings. Validity in this case is the fidelity with which three rating scales measure the efficiency of foresters. This is measured by correlating the final ratings obtained with the three rating scales with a criterion of success.



The criterion of success used is the average of the ratees' last three annual ratings. In instances where the ratee had been rated only twice, the average of the two ratings was used. Raters B and C made these three annual ratings on ratees one to seventeen.<sup>15</sup> One weakness of this criterion is that it is not an objective measure but a measure obtained with Form 3200, another rating scale.

Measures of validity cannot be expected to be extremely high even though a rating scale may be highly valid because of the weakness of the criterion and, also, all deficiencies in reliability directly lower the correlation of validity.

Correlations of final ratings made with each rating scale with a criterion of success consisting of the average of the ratees' last three annual ratings follow:

Rater and Men Rated:		Guide For Using Form 3200	Probst Service Report	Service Record
A	1 to 10	.65	.47	.65
B	1 to 10	.95	.92	.75
C	11 to 17	.92	.78	.40
D	11 to 17	.48	.90	.35
E	18 to 25	.72	.55	.43
Average		.75	.75	.53

The average correlations of .75 for both the Guide For Using Form 3200 and the Probst Service Report Rating Scales

15. Exceptions occur in instances where ratees have been transferred to either of the two Forests of which these raters are Forest Supervisors within the last three years.

are somewhat higher than the .53 correlation obtained with the Service Record Rating Scale.

On the whole the correlations made by raters B and C are higher than those of the other raters. This was to be expected as the criterion was largely made up of annual ratings which were made by them. Also the correlations made by raters B and C with the Guide For Using Form 3200 Rating Scale are especially high, .95 and .92, respectively. This can be accounted for in part by the fact that the criterion is an average of ratings obtained with a rating scale which has the same fundamental points and wordings as the Guide For Using Form 3200 Rating Scale.

It is quite significant that the correlation average obtained with the Probst Service Record Rating Scale of .75 is as high as that obtained with the Guide For Using Form 3200 Rating Scale. In this case the rating is correlated with a criterion which is also a rating but there is very little similarity between the two rating scales.

Although the Service Record Rating Scale is less valid than the other two rating scales it is not abnormally low under the circumstances. This lower validity may be due to weakness in the criterion.

## CHAPTER XIV

### SUMMARY

## CHAPTER XIV

## SUMMARY

Rating scales are measuring instruments designed to measure such physically unmeasurable qualities as initiative and cooperativeness which management considers important. They give management a basis for making many decisions relative to personnel problems such as promotions, salary changes, transfers, and employee training. Ratings also have a stimulating and educational effect upon the rater as well as the ratee.

The faultiness of human judgment in rating intangible personal qualities is recognized. There is the possibility of hidden bias, prejudice or favoritism. There are temperamental differences which lead some raters to rate most of their workers high, others to rate them low. Many problems of human memory and thinking occur, such as that of recalling fairly the performance of an employee during the full previous year, without allowing the excellent or the stupid work he may have done during the last week or month to dominate. Despite these difficulties employees are being rated in one way or another and must be rated in the ordinary conduct of business. Rating, therefore, is a practical and a basic problem. Rating scales attempt to solve this problem by organizing the process of measuring efficiency so that each employee may be as justly appraised as pos-

sible.

The first widespread use of rating scales was during the World War. Since that time many types and variations have been developed for use in business and government. Two of the most popular and practical types have been the graphic and the Probst rating scales. In the graphic type rating scale the range of ability for each quality appears on a horizontal line beside each quality while in the Probst Rating Scale the various degrees of performance appear in a list of about 100 observable actions.

A rating scale should measure the qualities that are considered important and eliminate the unimportant qualities. Therefore, in constructing a rating scale for foresters in the U. S. Forest Service an attempt was made to discover these important qualities. A questionnaire was sent to men in the field and authorities were consulted. The various traits and qualities were tabulated and six fundamental or essential qualities were found. These six qualities were experimentally evaluated and the results follow:

Intelligence -----	180 points
Personality -----	160 points
Character -----	180 points
Scientific Trend of Mind and	
Common Sense -----	160 points
Physical and Nervous Energy -----	144 points
Skill and Knowledge -----	144 points

These essential qualities were each defined in terms

of twelve to fifteen observable human actions in order to make them as objective as possible. These observable actions are the points in the rating scale that the rater considers in evaluating a ratee. Two methods of indicating a ratee's degree of perfection on each of the points listed were used. In one, called the Man-to-Man Record, the rater compares the ratee with a master scale consisting of three men representing the possible range of the ability. The ratee is checked on each trait as being most nearly like one of these three men. These steps are given values of twelve, six, and zero, respectively.

The other method of ratee evaluation, called the Service Record, lists five steps representing the possible range of performance. For example, a ratee is checked on the observable action, "His clothing is appropriate for the occasion," in one of the five ways as follows: nearly always, usually, about half of the time, occasionally, and hardly ever. These steps are given numerical values of twelve, nine, six, three, and zero, respectively.

In each of these rating scales the rater is not required to check any definite number of items. In fact he is asked to check items only when he is reasonably sure regarding the ratee's performance. This makes a special method of scoring necessary. The average score of the checked items is found in each of the six groups of traits.

Each unchecked item is given a score equal to the average item score in that group. All of the items are added together to get the score for each group and the total numerical score is the total of the six group scores. This special scoring is necessary in order to make the ratees' scores comparable.

Final numerical ratings are converted into one of five class ratings designed to contain ten, twenty, forty, twenty, and ten per cent of the total group in each class. These classes are represented by the letters A, B, C, D, and E. The boundaries of these classes were found statistically from a group of numerical ratings.

The two rating scales constructed, the Man-to-Man Record and the Service Record, were compared experimentally to see which was the more accurate. Ten raters rated one man. There was less variation with the Service Record Rating Scale and it was used in further experiments.

Service Rating Form 3200, which is a rating scale used in the Forest Service, the Probst Service Report, and the Service Record Rating Scales were compared experimentally in the School of Forestry to see which rating scale is the most consistent, or reliable. Professors rated a group of students in January and again in March. Average correlations of .90 and .89 showed that Form 3200 and the Probst Service Report, respectively, were quite reliable. The

Service Record Rating Scale was less reliable with a coefficient of .81. However, only about half of the items in this rating scale were checked as it calls for information regarding rates that professors do not have.

Attempts have been made to objectify Service Rating Form 3200 by converting it into a graphic rating scale called the Guide For Using Service Rating Form 3200.

An experiment was conducted in the Forest Service in order to compare three rating scales: Guide For Using Form 3200, Probst Service Report, and Service Record. The experiment was conducted normally through regular administrative channels.

The Guide For Using Form 3200 failed to segregate the employees into classes. The next to the highest class had the most ratings instead of the middle class. Both the Probst Service Report and the Service Record Rating Scales were satisfactory in this respect.

The Service Record Rating Scale was more reliable than the Guide For Using Form 3200 by twenty-one per cent with a coefficient of .98, as measured by the split-half technique. The Probst Service Report could not be measured by this method.

The validity of these three rating scales was measured by comparing the scores obtained with each rating scale with a criterion of success consisting of the average of



the ratees' last three annual ratings. The average coefficient of .75 for both the Guide For Using Form 3200 and the Probst Service Report Rating Scales showed them to be equally valid. However, the Guide For Using Form 3200 coefficient is probably positively affected by the similarity between the Guide For Using Form 3200 Rating Scale and Service Rating Form 3200, which was used in obtaining the scores which make up the criterion. The Service Record coefficient of .58 shows that it is somewhat lower in validity than the other rating scales but this may be due to the weakness of the criterion.

Results of the experiment conducted in the Forest Service show: The Guide For Using Form 3200 Rating Scale gives abnormally high final ratings and fails to segregate the ratees into their proper classes. Its reliability or consistency as a measuring instrument is questionable, and its validity is quite high. However, the value of this measure of validity is questioned because of the similarity of this rating scale and the one used in determining the criterion of success.

The Probst Service Report Rating Scale gives a normal distribution of final scores. It seems to be satisfactorily reliable, although other experiments should be made to test this, and it is equally as valid as the Guide For Using Form 3200 Rating Scale, which is noteworthy.

The Service Record Rating Scale also gives a normal distribution of final scores. It is highly reliable and it is satisfactory with respect to validity although it is lower in this respect than either of the other two rating scales.

Recommended Subjects for Additional Study.

1. A re-checking of the conclusions obtained in this paper using a larger number of samples.
2. Refining of the Service Record Rating Scale based upon experimental data gathered in this experiment.
3. Find the essential qualities of professional foresters in the United States Forest Service for each of the following groups: district rangers, members of supervisors' staffs, members of regional foresters' staffs, and research men.

## BIBLIOGRAPHY

BIBLIOGRAPHYBOOKS

- Baridon, Felix E., and Loomis, Earl H. Personnel problems. New York, McGraw-Hill Book Company, Inc., 1931.
- Bingham, W. V., and Freyd, Max. Procedures in employment psychology. New York, McGraw-Hill Book Company, Inc., 1926.
- Bossing, Nelson L. Progressive methods of teaching in secondary schools. San Francisco, Houghton Mifflin Co., 1935.
- Brumbaugh, H. C. Report on Probst Rating System by Multnomah County Civil Service Commission (mimeographed), December, 1938.
- Burt, Harold E. Principles of employment procedure. New York, Houghton Mifflin Company, 1926.
- Cleeton, G. U., and Mason, C. W. Executive ability, its discovery and development. Yellow Springs, O., Antioch Press, 1934.
- Commission of Inquiry on Public Service Personnel. Better government personnel. New York, McGraw-Hill Book Co., Inc., 1935.
- Commonwealth Club of California. Transactions. p. 323, 1921-22.
- Craig, David R., and Charters, W. W. Personal leadership in industry. New York, McGraw-Hill Book Company, Inc., 1925.
- Dashiell, John Frederick. Fundamentals of objective psychology. San Francisco, Houghton Mifflin Company, 1928.
- Dickinson, Z. C. Compensating industrial effort. New York, Ronald Press Company, 1937.
- Feldman, Herman. A personnel program for the Federal Civil Service; a report by the Personnel Classification Board. Washington, D. C., 1931.
- Ford, Adelbert. A scientific approach to labor problems. New York, McGraw-Hill Book Company, Inc., 1931.

- Garrett, Henry E. Statistics in psychology and education. Second Edition, New York, 1937.
- Graves, Henry S., and Guise, Cedric H. Forest education. New Haven, Yale University Press, 1932.
- Griffing, John B. Fundamentals of leadership, lessons one to five (Headquarters Ninth Corps Area, Office of the Educational Advisor, Presidio of San Francisco).
- Hackett, J. D. Labor management. New York, D. Appleton and Company, 1929.
- Illick, Joseph S. An outline of general forestry. New York, Barnes and Noble, Inc., 1936.
- Kelley, Evan W. What the forest service expects of forest school students. Mimeographed pamphlet, dated February 2, 1939.
- Kylie, H. R., Hieronymus, G. H., and Hall, A. G. CCC forestry. Washington, United States Government Printing Office, 1937.
- Kaird, Donald A. The psychology of selecting men. New York, McGraw-Hill Book Company, Inc., 1925.
- Lindberg, R. C. A report of Region 6 Advanced Forest Officers' Training Camp of 1935.
- Mayers, Lewis. The Federal Service. New York, D. Appleton and Company, 1922.
- Meriam, Lewis. Public personnel problems. Washington, The Brookings Institute, 1938.
- Meyers, Harry. Human engineering. New York, Harper and Brothers, 1932.
- Moon, Franklin, and Brown, Nelson Courtlandt. Elements for forestry. New York, John Wiley and Sons, Inc., third edition, 1937.
- Probst, J. B. Service ratings. Baltimore, The Lord Baltimore Press, 1931.
- Public Personnel Studies. What's wrong with service (efficiency) rating? Vol. 7, 1929.

- Schell, Erwin Haskell. The technique of executive control. New York, McGraw-Hill Book Company, Inc., 1930.
- Scott, Walter Dill, Clothier, Robert C., and Mathewson, Stanley B. Personnel management. Second edition, New York, McGraw-Hill Book Company, Inc., 1923.
- Tead, Ordway. The art of leadership. New York, McGraw-Hill Book Company, Inc., 1935.
- Tead, Ordway. Human nature and management. New York, McGraw-Hill Book Company, Inc., 1929.
- Tead, Ordway, and Metcalf, H. C. Personnel administration. Third edition, New York, McGraw-Hill Book Company, Inc., 1933.
- United States Civil Service Commission. Classification statutes. Washington, D. C., United States Government Printing Office, 1938.
- United States Civil Service Commission. Preparation of efficiency ratings. Departmental Circular No. 133, May 15, 1935.
- United States Department of Agriculture. Careers in forestry. Washington, Miscellaneous Publication No. 249, 1938.
- United States Department of Agriculture, Office of Personnel. Efficiency rating manual. Washington, D. C., 1939.
- Viteles, Morris S. Industrial psychology. New York, W. W. Norton and Company, Inc., 1932.
- Walters, J. E. Applied personnel administration. New York, John Wiley and Sons, Inc., 1931.
- Watkins, Gordon S. Labor management. New York, McGraw-Hill Book Company, Inc., 1928.
- White, Leonard D. Public administration. New York, The Macmillan Company, 1926.
- Wilmerding, Lucious, Jr. Government by merit. Commission of Inquiry on Public Service Personnel (New York), 1935.

MAGAZINES

- Bills, M. A. A method for classifying the jobs and rating the efficiency of clerical workers. The Journal of Personnel Research, V. 1, Nos. 8 and 9, p. 384-393, 1922-1923.
- Bingham, W. V. Halo: Its prevalence and nature in estimates of objective traits and in inferential trait judgments. Psychological Bulletin, V. 35, No. 9, p. 641-2, November, 1938.
- Blog, Leon. Does the Probst Rating System rate? National Municipal Review, V. 20, p. 581, 1931.
- Clarke, Walter V. Rating employees. The Journal of Personnel Research, V. 15, No. 3, p. 100-104, Sept. 1936.
- Coyle, Chalice Kelly; Answer to a grievance. Personnel Journal, V. 17, No. 5, p. 179-182, November, 1938.
- Freyd, M. Graphic rating scale. Journal of Educational Psychology. V. 14, p. 83-102, 1923.
- Hill, R. L. Efficiency ratings. Journal of Personnel Research, V. 15:330-2, March, 1937.
- Jones, E. S. Personality terms commonly used in recommendations. Journal of Personnel Research, V. 2, p. 421-430, 1924.
- Kingsbury, Forrest A. Analyzing ratings and training raters. Journal of Personnel Research, V. 1, p. 377-383, 1922-1923.
- Kingsbury, Forrest A. Making rating scales work. Journal of Personnel Research, V. 4, p. 1-6, 1925-1926.
- Knight, F. B. The effect of the acquaintance factor upon personal judgments. Journal of Educational Psychology, V. 14, p. 129-142.
- Koch, Elers. Technical requirements for a forester in the Federal Service. Journal of Forestry, V. 35, p. 803-9, Sept., 1937.
- Kornhauser, A. W. Reliability of average ratings. Journal of Personnel Research, V. 5, p. 309-317, 1926-1927.

- Kornhauser, A. W. What are rating scales good for? *Journal of Personnel Research*, V. 5, p. 189-193, 1926-1927.
- Mason, David T. The requirements for an education of a forester. *Journal of Forestry*, V. 35, p. 545-9, June, 1937.
- Ordway, Samuel H. Jr., and Laffon, John C. Approaches to the measurement and reward of effective work of individual government employees. *National Municipal*, V. 24, Supplement, p. 559, 1935.
- Paterson, Donald G. The Scott Company Graphic Rating Scale. *Journal of Personnel Research*, V. 1, Nos. 8 and 9, p. 361-376, 1922-1923.
- Ream, M. J. A statistical method for incomplete order of merit ratings. *Journal of Applied Psychology*, V. 5, p. 261-266, 1921.
- Riebold, R. J. In-service training for foresters. *Journal of Forestry*, V. 36, p. 146-8, Feb., 1938.
- Rugg, H. O. Is the rating of character possible? *Journal of Educational Psychology*, V. 12, p. 425-438 and p. 485-501, 1921; V. 13, p. 30-42, 81-93, 1922.
- Shen, E. The influence of friendship upon personal ratings. *Journal of Applied Psychology*, V. 9, p. 66-68, 1 1925.
- Thorndike, E. L. A constant error in psychological rating. *Journal of Applied Psychology*, V. 4, p. 25-29, 1920.
- Trow, William Clark. How shall teaching be evaluated? *Educational Administration and Supervision*, V. 20, p. 264-72, April, 1934.
- Watson, G. B. The measurement of the less tangible qualities. *Vocational Guidance Magazine*, V. 4, p. 281-289, 1926.
- Wohlenberg, E. T. F. Leadership in forestry and lumbering. *Journal of Forestry*, V. 31, p. 308-10, March, 1933.



## APPENDIX

TABLE V

Ratings Made in January by Four School of Forestry Professors of Students,  
with Three Rating Scales.

Ratees	Professor No. 1			Professor No. 2			Professor No. 3			Professor No. 4		
	Form	Probst	Service	Form	Probst	Service	Form	Probst	Service	Form	Probst	Service
	3200	Report	Record	3200	Report	Record	3200	Report	Record	3200	Report	Record
F 1	93	24	744	91	26	766	95	28	973	94	46	891
F 2	86	4	599	87	12	776	87	- 8	706	80	0	633
F 3	94	32	740	90	24	830	94	38	899	90	28	844
F 4	89	12	666	90	20	782	89	26	868	86	- 6	757
F 5	92	20	725	86	22	705	88	- 2	650	85	24	761
F 6	88	16	660	84	14	600	89	-14	824	77	-32	465
F 7	90	2	755	87	- 2	634	89	- 4	794	84	-10	577
F 8	77	-16	428	84	12	757	85	- 8	583	65	-22	584
F 9	94	38	753	90	34	854	94	32	826	91	28	850
F 10	93	46	747	91	36	887	94	24	879	94	38	864

TABLE VI

Ratings Made in March by Four School of Forestry Professors of Students,  
with Three Rating Scales.

Ratees	Professor No. 1			Professor No. 2			Professor No. 3			Professor No. 4		
	Form	Probst	Service	Form	Probst	Service	Form	Probst	Service	Form	Probst	Service
	3200	Report	Record	3200	Report	Record	3200	Report	Record	3200	Report	Record
F 1	92	52	751	96	36	853	93	22	853	93	42	920
F 2	85	0	555	90	26	707	90	10	374	87	4	879
F 3	93	52	846	96	28	936	93	38	894	95	40	939
F 4	89	30	681	94	20	879	93	18	751	90	12	897
F 5	91	42	751	92	22	837	91	6	618	91	20	933
F 6	88	16	540	91	22	790	87	-12	591	85	-12	653
F 7	90	20	683	92	10	726	92	6	702	87	8	623
F 8	80	-14	440	90	14	816	89	-10	612	81	-28	644
F 9	93	50	882	96	38	983	93	30	777	93	36	947
F 10	92	56	878	96	46	998	93	38	809	94	52	963

TABLE VII

Scores Made by Forest Service Raters Using  
 "Guide For Using Form 3200," "Probst Service Report,"  
 and "Service Record" Rating Scales

R a t e r s						
Ra- tees	A			B		
	Guide For F. 3200	Probst Report	Service Record	Guide For F. 3200	Probst Report	Service Record
1	90	18	713	88.5	10	645
2	92	58	765	93.5	76	902
3	88	- 4	619	88	-20	529
4	90	6	672	91	28	830
5	91	42	684	91	36	675
6	92	36	748	93	66	959
7	89	2	656	88	-10	681
8	87	-12	617	89	18	704
9	90	40	725	90	30	814
10	92	62	732	89	14	728
D			C			
11	91	32	729	92	52	695
12	91	34	655	91	42	711
13	89	14	663	90	34	718
14	91	22	829	90	46	763
15	91	4	594	88	10	599
16	90	14	739	89	12	736
17	84	-36	431	81	-52	395
F			E		G	
			Guide For Form 3200		Probst Report	
18	86	-12	420	87	-26	
19	83	-26	331	87	-34	
20	86	- 4	498	87	2	
21	87	-26	493	88	- 2	
22	83	6	544	87	- 8	
23	90	32	728	90	38	
24	88	16	655	86	14	
25	87	20	778	87	32	

TABLE VIII

Class Ratings Converted From Numerical Ratings Shown in  
Table VII by Forest Service Raters Using  
"Guide For Using Form 3200,"  
"Probst Service Report," and "Service Record" Rating Scales

R a t e r s						
Ra- tees	Guide For Form 3200	A		Guide For Form 3200	B	
		Probst Report	Service Record		Probst Report	Service Record
1	Very Good	C	C	Very Good	C	C
2	Excellent	B+	B	Excellent	A	A
3	Very Good	D	C	Very Good	D-	D
4	Very Good	C-	B	Very Good	C+	B
5	Very Good	B	C	Very Good	C+	C
6	Excellent	C+	B	Excellent	A	A
7	Very Good	D	C	Very Good	D	C
8	Very Good	D-	C	Very Good	C	C
9	Very Good	B	C	Very Good	C+	B
10	Excellent	A	C	Very Good	C	C
		D			C	
		Probst Report	Service Record		Probst Report	Service Record
11	Very Good	C+	C	Excellent	B+	C
12	Very Good	C+	C	Very Good	B	C
13	Very Good	C	C	Very Good	C+	C
14	Very Good	C	B	Very Good	B	B
15	Very Good	C-	D	Very Good	C-	D
16	Very Good	C	B	Very Good	C-	B
17	Good	E	E	Good	E	E
		F				
		Probst Report	Service Record			
18	Very Good	D-	E	Excellent is comparable to A; Very Good to B; Good to C; Fair to D; and Un- satisfactory is compar- able to E.		
19	Good	D-	E			
20	Very Good	D	E			
21	Very Good	D-	E			
22	Good	C-	D			
23	Very Good	C+	C			
24	Very Good	C	C			
25	Very Good	C	B			

All three of these rating scales have five final rating classes. The final scores on the "Guide For Form 3200" are directly comparable to the "Probst Service Report" and the "Service Record" Rating Scales with the above interpretations.

TABLE IX

Total Numerical Scores and Average Score Per Item Checked,  
Made by Forest Service Raters on the Odd and Even Numbered  
Questions Using "Guide For Using Form 3200" Rating Scale

(Used in making split-half correlations for reliability)

Rater A Rated Men 1 to 10

Rater D Rated Men 11 to 17

Rater F Rated Men 18 to 25

Ra- tees	Score Odd Items	No. Items	Average Score Per Item	Score Even Items	No. Items	Average Score Per Item
1	25	8	3.12	19	7	2.71
2	17	8	2.12	16	7	2.29
3	30	8	3.75	26	7	3.71
4	25	8	3.12	20	7	2.86
5	22	8	2.75	18	7	2.57
6	18	8	2.25	17	7	2.43
7	27	8	3.37	24	7	3.43
8	33	9	3.67	39	9	4.33
9	27	9	3.00	22	8	2.75
10	19	9	2.11	21	9	2.33
11	23	9	2.56	23	9	2.56
12	24	9	2.67	22	9	2.44
13	27	8	3.37	23	7	3.29
14	22	8	2.75	17	7	2.43
15	22	8	2.75	20	7	2.86
16	25	8	3.12	23	7	3.29
17	46	8	5.75	30	7	4.29
18	36	8	4.50	29	7	4.14
19	44	8	5.50	35	7	5.00
20	34	8	4.25	31	7	4.44
21	34	8	4.25	29	7	4.14
22	43	8	5.37	26	7	3.71
23	22	8	2.75	21	7	3.00
24	30	8	3.75	26	7	3.71
25	37	9	4.11	37	9	4.11

(Continued)

TABLE IX (Continued)

Rater B Rated Men 1 to 10 Rater C Rated Men 11 to 17						
Ra- tees	Score Odd Items	No. Items	Average Score Per Item	Score Even Items	No. Items	Average Score Per Item
1	31	8	3.87	22	7	3.14
2	16	8	2.00	12	7	1.71
3	32	8	4.00	26	7	3.71
4	22	8	2.75	17	7	2.43
5	21	8	2.62	18	7	2.57
6	15	8	1.87	13	7	1.86
7	31	8	3.87	25	7	3.57
8	28	9	3.11	26	8	3.25
9	27	9	3.00	23	8	3.82
10	28	9	3.11	26	8	3.25
11	16	9	1.78	23	9	2.56
12	21	9	2.33	27	9	3.00
13	24	8	3.00	22	7	3.14
14	22	8	2.75	21	7	3.00
15	17	5	3.40	17	4	4.25
16	24	6	4.00	19	6	3.17
17	30	5	6.00	24	4	6.00

TABLE X

Total Actual Numerical Scores and Average Score Per Item  
Checked, Made by Forest Service Raters on the Odd and Even  
Numbered Questions Using Service Record Rating Scale

(Used in making split-half correlations for reliability)

Rater A Rated Men 1 to 10  
Rater D Rated Men 11 to 17  
Rater F Rated Men 18 to 25

Ra- tees	Score Odd Items	No. Items	Average Score Per Item	Score Even Items	No. Items	Average Score Per Item
1	368	42	8.64	342	41	8.34
2	384	42	9.14	372	41	9.07
3	306	41	7.46	297	41	7.24
4	300	38	7.89	306	38	8.05
5	348	42	8.29	327	41	7.98
6	339	38	8.92	336	38	8.84
7	327	42	7.79	321	41	7.83
8	294	40	7.35	294	40	7.35
9	231	26	8.89	231	26	8.89
10	375	42	8.93	348	41	8.49
11	285	33	8.64	285	33	8.64
12	285	35	8.14	261	34	7.68
13	282	35	8.06	294	34	8.65
14	345	35	9.86	339	34	9.97
15	237	33	7.18	237	33	7.18
16	309	36	8.58	306	35	8.74
17	204	37	5.51	171	36	4.75
18	153	30	5.10	153	30	5.10
19	147	32	4.59	111	31	3.58
20	156	26	6.00	153	25	6.12
21	219	35	6.26	195	35	5.57
22	213	32	6.66	204	31	6.58
23	297	34	8.73	300	34	8.82
24	279	37	7.54	294	36	8.17
25	162	17	9.53	168	17	9.88

(Continued)



TABLE X (Continued)

Rater B Rated Men 1 to 10 Rater C Rated Men 11 to 17						
Ra- tees	Score Odd Items	No. Items	Average Score Per Item	Score Even Items	No. Items	Average Score Per Item
1	246	32	7.69	243	32	7.59
2	453	42	10.68	438	41	10.68
3	240	39	6.15	255	38	6.71
4	258	25	10.32	225	24	9.37
5	306	37	8.27	279	36	7.75
6	486	42	11.57	465	41	11.34
7	306	38	8.05	306	38	8.05
8	306	37	8.27	306	37	8.27
9	276	28	9.86	282	28	10.07
10	306	36	8.50	312	35	8.91
11	297	34	8.73	267	33	8.09
12	306	35	8.74	285	34	8.38
13	306	34	9.00	285	33	8.64
14	312	34	9.18	297	33	9.00
15	246	33	7.45	237	32	7.41
16	276	31	8.90	276	30	9.20
17	162	31	5.23	147	31	4.74

TABLE XI

Conversion Table for Changing Total Numerical Scores Into Final Ratings in "Form 3200" or "Guide for Using Form 3200," "Probst Service Report," and "Service Record" Rating Scales

---

Form 3200 or Guide For Using Service Rating Form 3200:

---

3 to 7	Excellent
8 to 13	Very Good
14 to 19	Good
20 to 24	Fair
25 to 30	Unsatisfactory

---

Probst Service Report (with a single rater):

---

60 to 80	A
40 to 60	B
4 to 40	C
-44 to 4	D
-80 to -44	E

---

Service Record:

---

836 to 1012	A
735 to 836	B
605 to 735	C
504 to 605	D
280 to 504	E

---

Form 3200—(April 1935)  
U. S. Civil Service Commission

RECTO

CLASSIFICATION SYMBOLS		
Service	Grade	Class

## SERVICE RATING FORM

(Read instructions on back of this form)

Check one:

Supervisory ..... ☐

Nonsupervisory ..... ☐

Name ..... Department .....

(Bureau) (Division) (Section) (Subsection)

On lines below  
mark employee:

✓ If neither strong  
nor weak point.

- If weak point.

+ If strong point.

1. Underline the elements which are especially important in the position.
2. Mark nonsupervisory employees on all elements except those in *italics*.
3. Mark supervisory employees on all elements.

In boxes below rate employee:

1 or 2 if Excellent.

3 or 4 if Very Good.

5 or 6 if Good.

7 or 8 if Fair.

9 or 10 if Unsatisfactory.

### I. QUALITY OF PERFORMANCE

- ..... (a) Acceptability of work; thoroughness
- ..... (b) General dependability; accuracy.
- ..... (c) Neatness and orderliness of work.
- ..... (d) Skill with which the important procedures, instruments, or machines are employed in performing his duties.
- ..... (e) *Effectiveness in getting good work done by his unit.*

Rating  
Officer

Reviewing  
Officer

### II. PRODUCTIVENESS

Base rating primarily on element (a), if known; otherwise on (b) and (c).

- ..... (a) Amount of work accomplished.
- ..... (b) Application of time, interest, and energy to duties; industry.
- ..... (c) Promptness in completing assignments; speed.
- ..... (d) .....
- (State any other elements of this class considered)
- ..... (e) *Effectiveness in securing adequate output from his unit.*

### III. QUALIFICATIONS SHOWN ON JOB

- ..... (a) Knowledge of duties and related information.
- ..... (b) Ability to learn and to profit from experience.
- ..... (c) Judgment, sense of proportion, common sense.
- ..... (d) Initiative and resourcefulness.
- ..... (e) Cooperativeness; ability to work with and for others.
- ..... (f) .....
- (State any other elements of this class considered)
- ..... (g) *Effectiveness in developing and training employees.*
- ..... (Custodial only) Ability to perform such physical work as the job requires.

Sum of ratings . . .

Rated by .....  
(Rating officer) (Date)

Report to employee .....

Reviewed by .....  
(Reviewing officer) (Date)

On the whole, do you consider  
the deportment and attitude  
of this employee toward his  
work to be satisfactory?

(Answer "Yes", "No", or "Fairly so")

Sum of Ratings	Report to Employee	Significance
3 - 7	Excellent.	Promotable within grade if below top salary.
8 - 13	Very Good.	Promotable within grade if below top salary.
14 - 19	Good.	No salary change if receiving middle salary or above; if below middle, promotable not beyond middle salary.*
20 - 24	Fair.	Reduce one step if above middle salary.*
25 - 30	Unsatisfactory.	Dismiss from present position.

\* For Cu-2 and Cu-3, the fourth salary rate will be considered the middle salary.

16-3831

(Back of Form 3200 same as back of Form 3201)

Form 3201- (April 1935)  
U. S. Civil Service Commission

**RECTO**

**SERVICE RATING FORM**

(Read instructions on back of this form)

CLASSIFICATION SYMBOLS		
Service	Grade	Class

Check one:

Supervisory ☐

Nonsupervisory ☐

Name ..... Department .....

(Bureau) (Division) (Section) (Subsection)

On lines below  
mark employee:  
✓ if neither strong  
nor weak point.  
- if weak point.  
+ if strong point.

1. Underline the elements which are especially important in the position.
2. Mark nonsupervisory employees on all elements except those in *italics*.
3. Mark supervisory employees on all elements.

In boxes below rate employee:  
1 or 2 if Excellent.  
3 or 4 if Very Good.  
5 or 6 if Good.  
7 or 8 if Fair.  
9 or 10 if Unsatisfactory.

**I. QUALITY OF PERFORMANCE**

- ..... (a) Thoroughness; adequacy of results.
- ..... (b) General dependability; accuracy of results.
- ..... (c) Technical skill with which the important procedures or instruments are employed in performing his duties.
- ..... (d) Original contributions to method or knowledge.
- ..... (e) *Effectiveness in getting good work done by his unit.*

Rating  
Officer

Reviewing  
Officer

**II. PRODUCTIVENESS**

- ..... (a) Amount of work accomplished.
- ..... (b) Application of energy, interest, and technical resources to duties; industry.
- ..... (c) Effectiveness in planning so as to utilize time to best advantage.
- ..... (d) Completing assignments; making progress on assigned projects.
- ..... (e) Composing adequate reports or other required writings.
- ..... (f) .....
- ..... (g) *Effectiveness in securing adequate output from his unit.*

**III. QUALIFICATIONS SHOWN ON JOB**

- ..... (a) Knowledge of particular field of work and of the fundamentals on which it is based.
- ..... (b) Analytical ability; constructive reasoning in the field of specialization.
- ..... (c) Scientific or professional attitude; fairness, freedom from bias.
- ..... (d) Judgment, sense of proportion, common sense.
- ..... (e) Initiative, resourcefulness; ability to grow.
- ..... (f) Cooperativeness; ability to work with and for others.
- ..... (g) .....
- ..... (h) *Effectiveness in developing and training employees.*

Sum of ratings . . .

Rated by .....  
(Rating officer) (Date)

**Report to employee**

Reviewed by .....  
(Reviewing officer) (Date)

On the whole, do you consider  
the department and attitude  
of this employee toward his  
work to be satisfactory? . . .

(Answer "Yes", "No", or "Fairly so")

Sum of Ratings	Report to Employee	Significance
3 - 7	Excellent.	Promotable within grade if below top salary.
8 - 13	Very Good.	Promotable within grade if below top salary.
14 - 19	Good.	No salary change if receiving middle salary or above; if below middle, promotable not beyond middle salary.
20 - 24	Fair.	Reduce one step if above middle salary.
25 - 30	Unsatisfactory.	Dismiss from present position.

(See next page for back of Forms 3200 and 3201)

CIVIL SERVICE FORMS 3200 and 3201 - VERSO

## CONDUCT REPORT

(This space is to be used in case the question on the face of the sheet, regarding the employee's deportment and attitude, has been answered "No" or "Fairly so." In such a case give here a full statement of the particulars in which the employee's conduct has been unsatisfactory.)

-----  
(Rating officer)

## INSTRUCTIONS TO RATING OFFICERS

1. Compare the qualifications and performance of each employee, as demonstrated by his work, with the actual needs of the position, considering the conditions under which the work must be done. Beginning with the lowest grade (CAF-1, P-1, SP-1, Cu-1, or CM-1), rate each series of classes (such as Junior Stenographer, CAF-2, Senior Stenographer, CAF-3, etc.) as a separate group. Keep in mind reasonable standards of performance for the various grades. The same rating standards should be applied to all competing employees in the same grade, irrespective of the fact that some may be receiving compensation at the minimum pay rate of the grade and others at higher rates.

2. The elements (a), (b), (c), etc., listed under each title (I, II, III) are not of equal importance. Underline the elements which are especially important in the position.

3. If the performance of an employee is neither strong nor weak with respect to an element, put a check mark (✓) on the line at the left of the element; if weak, a minus (−); if strong, a plus (+).

Differentiate carefully among the several elements. Extreme care should be taken to avoid basing all marks on some one strong or weak characteristic of the employee. A person who deserves a plus or minus on one element does not necessarily merit the same mark on all elements.

4. If in your judgment the employee is excellent on "I. Quality of Performance", indicate this by 1 or 2 in the box at the right;

if very good ..... by 3 or 4

if fair ..... by 7 or 8

if good ..... by 5 or 6

if unsatisfactory ..... by 9 or 10

Indicate your ratings on titles II and III in the same manner.

The numerical ratings on the titles (I, II, and III) are not derived by a mechanical summary of the element marks, but depend on the best judgment of the rating officer as to how well the employee meets the broader requirements of the position. This judgment is assisted by the element marks, but is not rigidly determined by them. These marks insure that the employee's performance on the elements which affect Quality of Performance, Productiveness, and Qualifications Shown on the Job will be considered.

5. The rating to be reported to the employee by the board of review is the adjective corresponding to "Sum of ratings" on I, II, and III, as given in the table at the bottom of the rating form.

6. Marks and ratings should first be made lightly with pencil. After all your employees have been rated, compare and consider the marks and ratings assigned to the various employees in the same classes, make any necessary alterations, and indicate the marks and ratings in black ink.

7. The name of an employee rated on a service of less than 90 days should be followed by the notation, "Less than 90 days."

8. The question on deportment should be answered "Yes", "No", or "Fairly so." If the answer is "No" or "Fairly so", it should be explained in the space provided above. Rating officers should not allow unsatisfactory conduct to influence marks or ratings, except as it may actually affect an employee's performance on some specific element or elements.

9. Complete the ratings promptly. Submit the signed and dated rating forms to the reviewing officer.

## INSTRUCTIONS TO REVIEWING OFFICERS

1. Compare the marks and ratings assigned by the different rating officers under your supervision, noting such corrections as may be necessary to secure reasonable uniformity of standards and accuracy in the marks and in the ratings.

2. Make corrections with red ink, but do not cross out or erase the marks or ratings made by the rating officer. Before any marks or ratings are changed discuss them with the rating officer.

3. Submit the signed and dated rating forms to the board of review promptly.

(Permission to use granted by UNITED STATES CIVIL SERVICE COMMISSION)

GUIDE FOR USING SERVICE RATING FORM #3200

I. QUALITY OF PERFORMANCE

Score

(a) Acceptability of work; thoroughness

1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9

Eager to accept  
work; very thorough

Accepts work;  
thorough

Rarely seeks work;  
occasionally  
lacks thoroughness

Never seeks work;  
lacks thorough-  
ness

(b) General dependability; accuracy

1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9

Very dependable;  
consistently  
accurate

Usually depend-  
able and ac-  
curate

Not dependable;  
frequently makes  
mistakes

Seldom does ac-  
curate work;  
always requires  
checking

(c) Neatness and orderliness of work

1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9

Work consistently  
neat and orderly

Work satisfac-  
tory but occa-  
sionally not  
orderly

Work frequently un-  
organized and not  
neat

Work unsatis-  
factory

Score	(d) Skill with which the important procedures, instruments, or machines are employed in performing his duties							
	'	'	'	'	'	'	'	'
	2			4		6		8

Uses excellent procedure      Good procedure      Fair procedure      Poor procedure used

	(e) Effectiveness in getting good work done by his unit							
	'	'	'	'	'	'	'	'
	2			4		6		8

Well organized; accurate work      Well organized; usually accurate      Fairly well organized; occasional errors      Poorly organized with frequent errors

Total      Average for I.

## II. PRODUCTIVENESS

Score	(a) Amount of work accomplished							
	'	'	'	'	'	'	'	'
	2			4		6		8
	Above the average		Average		Below average		Unsatisfactory	

Score

(b) Application of time, interest and energy  
to duties; industry

1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	1

Always looking for  
work

Average, moder-  
ately energetic

Not anxious to work Lazy

(c) Promptness in completing assignments; speed

1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	1

Always works rapidly  
and efficiently

Works diligently  
at a moderate  
rate

Works slowly  
Unsatisfactory

\*(d) Effectiveness in securing adequate output  
from his unit

1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	1

Above the average

Average

Below average

Unsatisfactory

Average for II.

Total



### III. QUALIFICATIONS SHOWN ON JOB

## Score

(a) Knowledge of duties and related information

Knowledge of duties clearly demonstrated in work	Average knowledge of duties and related information	Lack of knowledge of duties apparent	Marked inefficiency due to lack of knowledge of duties
--------------------------------------------------	-----------------------------------------------------	--------------------------------------	--------------------------------------------------------

(b) Ability to learn and to profit from experience

[illegible]

Appreciates criticism; makes fullest use of criticism and experience	Usually appreciates criticism and uses it to advantage	Occasionally resents criticism; does not use it	Resents criticism and does not profit by it
----------------------------------------------------------------------	--------------------------------------------------------	-------------------------------------------------	---------------------------------------------

(c) Judgment, sense of proportion, common sense

Clear and exact thinking; excellent judgment	Occasionally somewhat confused; good judgment	Frequently becomes confused; poor judgment	Thinking not logical; unsatisfactory
----------------------------------------------	-----------------------------------------------	--------------------------------------------	--------------------------------------

Score

(d) Initiative and resourcefulness

1	2	3	4	5	6	7	8	9
	2		4		6		8	

Initiative highly developed

Plans work with Below average; must Practically no moderate amount frequently be told initiative of supervision what to do

(e) Cooperativeness--ability to work with and for others

1	2	3	4	5	6	7	8	9
	2		4		6		8	

Always agreeable and cooperative

Occasionally disagreeable and difficult to co-operate

Frequently disagreeable and not cooperative

Unable to co-operate and get along with fellow workers

\*(f) Effectiveness in developing and training employees

1	2	3	4	5	6	7	8	9
	2		4		6		8	

Excellent

Very good

Good or fair

Unsatisfactory

Total

\*Considered only for supervisory employees.

Average for III.

COPYRIGHT, 1932, BY J. B. PROBST  
IN THE UNITED STATES  
AND CANADA  
ALL RIGHTS RESERVED

**RECTO**  
**THE PROBST SERVICE REPORT**

GENERAL FORM

FOR APPRAISING THE SERVICE VALUE OF  
EMPLOYEES, SUPERVISORS OR OFFICERS,  
OTHER THAN THOSE IN THE POLICE, FIRE,  
LABOR, OR EDUCATIONAL SERVICES

(GENERAL FORM)

The facts and judgments recorded on this sheet  
are evaluated by a scientifically constructed pro-  
cess and formula, producing a letter rating which  
takes into account all the checked items.

RATING
--------

FOR THE SIX-MONTH PERIOD

ENDING \_\_\_\_\_

NAME OF EMPLOYEE

TITLE

DISTRICT, DIVISION, OR STATION

**INSTRUCTIONS**

1. On this form you are to report the service value of the employee mentioned above. The report should be for the six-month period shown hereon, unless otherwise indicated.
2. In addition to the blanks to be filled in on this side of the sheet, you should check (with an X) all those items on the other side that you can find which will properly fit or describe this employee. Do not guess; if you are not reasonably sure that the employee possesses the trait or quality indicated by a certain item, do not check that item at all. It is not necessary to check any given number of items. You may be able to check 25 or more for one employee and have difficulty in finding more than a dozen or so to describe properly some other employee. Make your X's small; keep them inside the little squares. Do not change the wording of any item.
3. This sheet should be checked by three supervisors, wherever possible. Each supervisor should select one of the three check columns in which to make his X marks, and should keep all his marks within that same column on both sides of the sheet. The supervisor who is lowest in rank or authority should be the first to check the sheet; then the next higher (or equal) in rank should check; and the one in highest authority should check last. (See direction booklet.)
4. Some items, such as "Good technical training for the work", "Good headwork in sudden emergencies", and a few others, should be considered only if they are deemed essential or desirable for the particular position.

How many days was this employee absent during this  
period—(Do not include absence for injury in line of duty  
or absence on regular vacation)

(a) For sickness, with pay \_\_\_\_\_ days

(b) For sickness, without pay \_\_\_\_\_ days

(c) For personal reasons, with or without pay \_\_\_\_\_ days

(d) How many times absent without leave \_\_\_\_\_

(e) (If there was any other absence, or suspension, or loss of vacation, bonus,  
or merits, or other penalty; explain briefly here.) \_\_\_\_\_

Check only one item in each of the following two boxes. Consider not only the punctuality of the employee in reporting for work, but also his promptness in answering calls, keeping appointments, submitting reports, and doing specially assigned work.

Check Columns			
	1	2	3
BOX 1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			Nearly always late
			Usually late
			Often late (about half the time)
			Usually punctual
			Never, or hardly ever, late

BOX 2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			Nearly always quits ahead of time
			Usually quits ahead of time
			Often quits ahead of time
			Watches clock too much near quitting time
			Seldom quits ahead of time
			Never quits ahead of time

Identification key to reporting officer		
Column 1	Column 2	Column 3

--

THE PROBST SERVICE REPORT - VERSO

(GENERAL FORM)

**DIRECTIONS:** Place an X mark next to each of the items on this page which you know from your own knowledge will describe or fit this employee. Do not guess; check only if you are reasonably certain.

Check Columns			
1	2	3	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Lazy
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Slow moving
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Quick and active
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Too old for the work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Minor physical defects
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Serious physical defects
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Indifferent; not interested
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Talks too much
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Too blunt or outspoken
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Too much self-importance
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Good team worker
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Not a good team worker
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Resents criticism or suggestions
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Antagonizes when dealing with others
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Might often be more considerate
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Usually pleasant and cheerful
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Always* courteous
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Cranky disposition
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Often seems dissatisfied
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Often grumbling or complaining
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Uses poor judgment
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Might often use better judgment
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Generally uses good judgment
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Always* uses good judgment
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Does not do his (her) share of work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Generally looks for the easy work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Must generally be told what to do
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Work often slightly behind
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Often needs prodding
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Work always* up to date
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Turns out unusually large amount of work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Steady worker most of the time
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Always* busy at work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Does not accept responsibility
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Accepts responsibility
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Does not always obey orders willingly
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Visits too much with others
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Needs considerable supervision
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Works well without supervision
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Fine self-control; seldom or never loses temper
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Loses temper easily
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Easily rattled or upset
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Lacks self-confidence
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Too easy-going
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Learns new work slowly
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Learns new work easily
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Understands instructions readily
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	A willing worker at all times
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Takes unusual interest in the work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Might be more orderly
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Very orderly and systematic
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Often forgetful
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Often does careless work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Makes many mistakes
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Usually accurate
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Hardly ever makes a mistake
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Accurate but very deliberate
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Is highly expert in own work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Not generally reliable or dependable
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Usually reliable and dependable
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Always* reliable and dependable

Check one item only, if any.

Check one item only, if any.

Check one item only, if any.

Check Columns			
1	2	3	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Active and strong
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Active but not strong
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Usually careless of personal appearance
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Usually neat personal appearance
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Has a pleasing voice and manner
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Very tactful in dealing with the public
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Poor technical training for the work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Good technical training for the work
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Poor head work in sudden emergencies
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Good head work in sudden emergencies
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Often assigned to other important positions**
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Often assigned to fill a higher position**
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sometimes goes on a "tear"
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Drink is one of principal failings
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Willing worker, but is not a leader
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Does not plan or lay out work effectively
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Plans work well but lacks snap in getting it done
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Unusual ability in planning and laying out work; good organizer
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Makes quick and accurate decisions
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Too lenient in maintaining discipline
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Maintains good discipline
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Lacks decision
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Exceptionally skillful in handling difficult situations
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Makes poor sales talk
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Makes good sales talk
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Always tries to please
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Unusually gracious toward customers
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Often not attentive to customers
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Uses good English
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sales volume among the best
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Skillful in overcoming objections
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sales volume below average
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Voice too loud, harsh, or high pitched
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Keeps up fine display of goods
In the following spaces you may add such other items of your own as you believe will further describe this employee.			
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
REMARKS			

Use these items only where they apply, especially for supervisory positions.

Use these items only where they apply, especially for sales clerks and others who deal with customers.

\*Meaning "with rare exceptions"

\*\*Temporary assignments

(Reproduced by permission of copyright owner)