

A CONSTANT ERROR IN PSYCHOLOGICAL RATINGS

By EDWARD L. THORNDIKE, Teachers College, Columbia University



In a study made in 1915 of employees of two large industrial corporations, it appeared that the estimates of the same man in a number of different traits such as intelligence, industry, technical skill, reliability, etc., etc., were very highly correlated and very evenly correlated. It consequently appeared probable that those giving the ratings were unable to analyze out these different aspects of the person's nature and achievement and rate each in independence of the others. Their ratings were apparently affected by a marked tendency to think of the person in general as rather good or rather inferior and to color the judgments of the qualities by this general feeling. This same constant error toward suffusing ratings of special features with a halo belonging to the individual as a whole appeared in the ratings of officers made by their superiors in the army.

The official rating plan devised by Walter Dill Scott called for separate ratings for Physical Qualities, Intelligence, Leadership and Personal Qualities (i. e. Character). The instructions very emphatically required each of these four to be estimated independently of the others, as appears from the directions quoted below. Yet the correlations of the Intelligence rating with the ratings for Physique, Leadership and Character made by a very conscientious officer in the case of 137 aviation cadets whose work he, as flight commander, supervised, were .51, .58 and .64 respectively. These are all higher than reality, plus the attenuation due to erroneous judgments, could well give, especially within the restricted range of the commissioned-officer group. They are also too much alike. In reality Intelligence and Character or Intelligence and Leadership should give about three times as close a correlation as Intelligence and Physique.

" How to Make the Scale.

3. Make a list of about a dozen officers of your own rank and not above the average age of officers of this

rank. They should be men with whom you have served or with whom you are well acquainted. Include officers whose qualifications are poor or mediocre as well as those who are highly efficient. This list serves merely as a convenient reservoir of names; the names actually used in the scale may include others.

4. Look over your list from the viewpoint of *Physical Qualities* only. Disregard every characteristic of each officer except the way in which he impresses his men by his physique, bearing, neatness, voice, energy and endurance. Select that officer who surpasses all the others in this qualification and enter his name on the line marked highest under *Physical Qualities*. Now select the one who most conspicuously lacks these qualities and enter his name on the line marked lowest. Select the officer who seems about half way between the two previously selected and who represents about the general average in physical qualities; enter his name on the line marked middle. Select the officer who is half way between the middle and the highest; enter his name on the line marked high. Select the one who ranks half way between middle and lowest; enter his name on the line marked low.

5. In the same manner make out scales for each of the other four sections (*Intelligence, Leadership, Personal Qualities and General Value to the Service*).
How to Use the Scale.

6. Rate your subordinate for *Physical Qualities* first. Consider how he impresses his men by his physique, bearing, neatness, voice, energy and endurance. Compare him with each of the five officers in section I of the Rating Scale, and give him the number of points following the name of the officer he most nearly equals. If he falls between two officers in the Scale give him a number accordingly (e. g. if between Low and Middle give him 7, 7½ or 8).

7. Rate the subordinate in a corresponding manner for each of the other four essential qualifications. Under III (*Leadership*) and V (*General Value to the Service*) consider which officer he will most nearly equal after equivalent experience.

* * * * *

Points for Special Attention.

9. In making or using any section of the scale, consider only the qualification it covers, totally disregarding all the others.

I. Physical Qualities.

Physique, bearing, neatness, voice, energy and endurance. (Consider how he impresses his men in the above respects.)

II. Intelligence.

Accuracy, ease in learning, ability to grasp quickly the point of view of commanding officer, to issue clear and intelligent orders, to estimate a new situation, and to arrive at a sensible decision in a crisis.

III. Leadership.

Initiative, force, self reliance, decisiveness, tact, ability to inspire men and to command their obedience, loyalty and co-operation.

IV. Personal Qualities.

Industry, dependability, loyalty, readiness to shoulder responsibility for his own acts, freedom from conceit and selfishness, readiness and ability to co-operate.

V. General Value to the Service.

His professional knowledge, skill and experience; success as an administrator and instructor; ability to get results."

The same effect appears in the ratings given by other officers. The correlations are too high and too even. For example, for the three raters next studied the average correlation for physique with intelligence is .31; for physique with leadership, .39; and for physique with character, .28.

The same constant error appears in the correlation of the total Scott rating with a rating for technical ability as a flyer in the case of aviation officers. It is known from abundant evidence that technical ability as a flyer is a rather highly specialized quality.¹ Considering the restricted range of the aviation cadets, the correlation between general ability for officer work and technical ability as a flyer could hardly be above .40, without any attenuation. As attenuated by the imperfections of the rater's knowledge of both, it could hardly be above .25. Yet the correlations for the eight raters studied in this respect are .74, .85, .52, .91, .63, .72, .47 and .53, an average of .67. Obviously a halo of general merit is extended to influence the rating for the special ability, or vice versa.

Mr. Knight of Teachers College has studied this same effect in the case of 129 teachers rated by their superior officer

¹ See "The Selection of Military Aviators: Mental and Moral Qualities," *U. S. Air Service*, June 1919.

for certain qualities on the Boyce score card. The ratings in question were official and were used to determine salaries and promotions. General merit as a teacher has correlations of .68 with intellect, .79 with power in discipline, and .63 with voice. It is clear that the rating of a teacher's voice must have been influenced by the general impression of her ability. Voice correlates .50 with "Interest in Community Affairs," and .63 with intelligence!

The correlations reported in the original study by Boyce show this same effect. General merit as a teacher is reported to correlate as follows:

With General Appearance.....	47
" Health	56
" Voice	53
" Intellect	62
" Initiative and self reliance	77
" Adaptability and resourcefulness	80
" Accuracy	74
" Industry	69
" Enthusiasm	71
" Integrity and sincerity	63
" Self control	66
" Promptness	66
" Tact	69
" Sense of Justice.....	61
" Academic preparation	41
" Professional preparation	38
etc. etc.	

(The last is the lowest of the forty-five correlations reported.)

In the cases so far the correlations are a resultant of (1) the real facts, (2) the constant error of the "halo," as we may call it, and (3) the reverse error of attenuation due to chance inaccuracies in the ratings. In certain further work by Mr. Knight the correlations are freed from the last influence, by being based on the composite rating of two groups, each of a number of teachers who knew the individuals to be rated fairly well. The self-correlations of the ratings by one such a group with ratings for the same trait by the other group are over .90. The correlations for general ability as a teacher with intellect and with ability to discipline are about .95 and .80! The correlation of intelligence and ability to discipline is about .80! The correlations of a standard test of intelligence with general ability as a teacher and with ability to discipline are, for the individuals in question, not over .3.

The writer has become convinced that even a very capable foreman, employer, teacher, or department head is unable to treat an individual as a compound of separate qualities and

to assign a magnitude to each of these in independence of the others. The magnitude of the constant error of the halo, as we have called it, also seems surprisingly large, though we lack objective criteria by which to determine its exact size. As a consequence science seems to demand that, in all work on ratings for qualities the observer should report the evidence, not a rating, and the rating should be given on the evidence to each quality separately without knowledge of the evidence concerning any other quality in the same individual.