

MEASURING THE ABILITY TO JUDGE POETRY

MELVIN G. RIGG
Oklahoma A. and M. College

The measurement of aesthetic discrimination is of considerable importance to psychology, since it must precede any attempt to study the relationship of the aesthetic traits to other aspects of personality. Aesthetic judgment in the field of poetry may be measured by means of a test in which passages of accepted merit are to be compared with inferior parodies.

Development of the Test

The Rigg Poetry Test began with 104 items, each consisting of two passages, one taken from the work of a recognized poet, and the other a parody of it written especially for the test. An example will be given:

A She walks in beauty like the night
Of cloudless climes and starry skies;
And all that's best of dark and bright
Meets in her aspect and her eyes.

B My love is very beautiful, as beautiful
as night;
Like starry skies are her two eyes,
a rich and lovely sight.

In this item the A passage is from Byron; the B passage is the parody. This is one of the easier discriminations in the test. In the other items either the original or the parody may be printed first. The person taking the test is asked to select the passage which he regards as the better poetry.

The 104 original items were arranged in two experimental forms, Form A and Form B. On the basis of several criteria 24 of the items were eliminated. The remaining 80 items were divided into two forms, C and D, of 40 items each. The score on each form is the number right, 40 being perfect.

The Validity of the Test

Since in each item the passage written by the recognized poet is considered better than the parody written by a mere psychologist (who moreover was trying to produce something inferior), the validity of the test is based upon the verdict of history. However, as a further check the items were submitted to a group of "Experts," who were asked to give their opinions unbiased by considerations of authorship. This group of experts consisted of 47 college professors and 3 poets. Of the former, 43 were professors of English and 4 were professors of foreign literature. The items were finally passed in review by a panel of eight experts, who discussed in detail all the cases in which the discrimination is close. As a result of this technique the writer is confident that in each case the original is to be considered superior to the parody by the weight of contemporary expert opinion.

The Reliability of the Test

The reliability of the test was increased considerably during the early experimental stage by the method of internal consistency, since items contributing to unreliability were eliminated.

The reliabilities reported below are based upon correlations between Form C and Form D, given on separate days. The reliabilities of both forms together are estimated by means of the Spearman-Brown prophecy formula.

Group	Reliability Coefficients		
	N	rCD	Reliability C+D
High School	402	.72	.84
College Undergraduates	286	.75	.86
Total Adult	342	.82	.90

(The Total Adult group includes the college undergraduates, the experts, and a few other cases, mostly college graduates.)

Group	Tentative Norms for the Test		
	N	Medians	
		Form C	Form D
High School	402	20.4	21.3
College Undergraduates	286	25.5	26.1
Experts	50	36.7	35.5

Discussion of the Results

Although most of the experts made scores very near the top, in a few cases the scores were surprisingly low. There is on the whole much overlapping among the three groups.

Overlapping of the Three Groups

	Form C	Form D
No. of experts worse than the upper quartile of college undergraduates	4	2
No. of experts worse than the upper quartile of high school students	1	2
No. of college undergraduates better than the lower quartile of the experts	22	24
No. of high school students better than the lower quartile of the experts	8	17

One fourth of the high school students are better than the average college undergraduate. All these facts indicate that the ability to judge poetry is only partly dependent upon training.

The Use of the Test

In addition to its more obvious use as a means of appraisal in literature classes, the test may be employed as a teaching device. After the papers are scored, they can be returned to the pupils for a discussion as to why the various parodies are inferior.

But more important from a scientific point of view is the use of the test as a research tool in many problems of psychological aesthetics. One investigation based upon this test has already been completed, and indicates that the ability to judge poetry is not the same as the ability to judge music.¹

¹ Rigg, M., "The Relationship between Discrimination in Music and Discrimination in Poetry," *Journal of Educational Psychology*, February, 1937, pp. 149-152.