

The Catholic University of America

A Factor Analysis of Music Tests

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF ARTS AND SCIENCES OF THE CATHOLIC UNIVERSITY
OF AMERICA IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

by

LIBBIE B. BOWER, M.A.



THE CATHOLIC UNIVERSITY OF AMERICA PRESS
WASHINGTON, D. C.

1945

CONTENTS

	PAGE
INTRODUCTION	vii
CHAPTER	
I. REVIEW OF LITERATURE	1
II. PROCEDURE OF EXPERIMENT	9
III. CONSTRUCTION AND ADMINISTRATION OF THE MUSICAL MOODS TEST	17
IV. ANALYSIS OF AUDIOMETER TESTS	26
V. STATISTICAL DATA	31
SUMMARY AND CONCLUSIONS	36
APPENDIX I. Table I—Intercorrelations of 15 Variables	38
APPENDIX II. Table II—Factor Loadings Before Rotation and Factor Loadings After 4 Rotations	39
APPENDIX III. Table III—Hierarchy of Factors I, II, III ..	39
APPENDIX IV. Musical Moods Test	40
APPENDIX V. Questionnaire for Music Tests	41
BIBLIOGRAPHY	42
INDEX OF NAMES	45

INTRODUCTION

The development of standardized tests in the field of music has provided investigators with objective measuring instruments which may help to eliminate much of the subjective element in the evaluation of pupil ability. However, it is well known that the reliabilities of these tests are low and that they cannot compare in excellence with the best intelligence tests. Moreover, they are so few in number that they can at best define only a small part of the music field.

The question of whether aesthetic experience can be approached in the same way that scientific material is measured is discussed by Mainwaring.¹ He argues that the empirically and mathematically educed data provide material from which may be constructed something on which to build, no matter how nebulous this material may be.

The present study is written from the point of view that the musical ability of an individual, trained as well as untrained, may be partly explored in terms of three basic approaches, i.e., the sensory (auditory sensitivity), the perceptual (aesthetic appreciation), and the affective (musical feeling and understanding).

An exploratory investigation using tests within these three areas was conducted. The data was factorized with Thurstone's centroid technique. Other variables such as the Meier Art Test, a questionnaire and a rating scale, teachers' grades in music, intelligence quotients, and audiometer tests on one fourth of the subjects were further checks on the relationship of this complex ability, musical ability.

¹J. Mainwaring, "An Examination of the Value of the Empirical Approach to Aesthetics," *British Journal of Psychology*, XXXII (1941), p. 115.

CHAPTER I

REVIEW OF LITERATURE

A brief survey of the history of psychological studies on musicality show that different investigators differ as to the number of factors as well as to the importance they attach to one or more of these factors, in their estimation of musicality. Schoen² states the position of several psychologists in the early part of the 20th century. Stumpf had four criteria for musicality: (1) to sing a given tone correctly, (2) to discriminate higher and lower tones, (3) to judge whether one or two tones are present in a clang, and (4) to discriminate between degrees of consonance as to pleasantness. Wundt held that musically gifted persons were those who possessed sufficient power of musical hearing and memory to enable them to retain and recognize intervals at least for the duration of an experiment. Révész did an exhaustive longitudinal study on a child prodigy. He found that his subject, when given a series of tests on his auditory and musical powers, could instantly name any tune or chord played for him, and sing any desired note. He recognized every interval and gave its musical name. He named and analyzed the most complicated chords and dischords, and he could transpose a composition into any key faultlessly.

During the past twenty-five years, a large amount of work has been done and many studies have been conducted in applying to music the techniques of testing and measurement. Seashore,³ the outstanding American name in the psychology of musical talent, collated many of the earlier studies breaking down the undifferentiated concepts of musical talent into various aspects and devised an inventory which he termed "The barest skeletal outline of a musical mind."

² M. Schoen, *The Psychology of Music* (New York: The Ronald Press Co., 1940), 152-156.

³ C. E. Seashore, *The Psychology of Musical Talent* (New York: Silver Burdett Co., 1919), 288 pp.

One of these aspects, the hearing of tones, was constructed as a scientific tool for measuring auditory sensitivity and discrimination. This⁴ consists of six tests which were published in 1919 and revised in 1939. His work pioneered the way for other investigators whose efforts were directed toward the construction of objective tests which dealt with these as well as other factors. Some of these tests were standardized and were used with the revised Seashore battery⁵ in this study.

Burt,⁶ in 1925, on factorizing tested abilities among school children and students, reported evidence of a group factor for music, i.e., a special and distinct highly complex musical capacity. He concluded that a special or group factor for musical ability existed over and above the general factor for intelligence. As a result of his study, Burt strongly urged that in the experimental procedure, tests of musical ability should embrace the more complex and concrete processes such as are involved in ordinary listening, in order to deal with the essential aesthetic elements involved in musical appreciation, rather than limit the tests to those of auditory discrimination to which the tests of Seashore and others are chiefly confined. He points out that these tests rest on principles which would condemn half of the present day compositions and that they are entirely out of keeping with the style of music to which the younger generation has become accustomed. A few years later, Burt⁷ used tests for each of these aspects of music: appreciation of melody, of harmony, of rhythm, of chord progressions, of cadences, detecting appropriate or inappropriate instances of each, noting faults of execution, assigning the best titles, drawing up the best critical comments, programmes, and emotional descrip-

⁴ C. E. Seashore, *Measures of Musical Talent* (Chicago, Ill.: C. H. Stoelting Co., 1919), 16 pp.

⁵ J. G. Saetveit, Don Lewis, C. E. Seashore, *Revision of the Seashore Measures of Musical Talents* (Iowa City, Iowa: The University of Iowa Press, 1940), 7-35.

⁶ C. Burt, *Report of the Consultative Committee on Psychological Tests of Educable Capacity*, Board of Education (1925), H. M. Stationery Office, p. 20.

⁷ *Ibid.*, *The Psychology of Music*, Gresham Lectures, Unpublished, fide H. D. Wing, "A Factional Study of Musical Tests," *The British Journal of Psychology*, XXXI (1941), 342, 343.

tions, ranking tunes (either heard or referred to by title) according to relative merit. In spite of the fact that the preliminary results were imperfect, as the reliability coefficient was low (.60 to .75), three things were demonstrated: (1) the importance of the theoretical principle on which musical tests should be constructed, (2) the superiority of testing more complex and more natural types of listening, (3) the value of assessing the individual's general capacity for musical appreciation, as well as the particular type of musical appreciation in which he is weak or strong.

Spearman,⁸ in 1927, using Seashore's tests of pitch, loudness, and rhythm, brought forward reasons for doubting a group factor for music as well as other group factors. He said that no analogous broad factor has presented itself. Spearman⁹ said "It might have been expected in the sphere of music where not only innate instinct but also environmental encouragement are incomparably more favorable for some individuals than for others." He found that the abilities to appreciate relations of pitch, loudness, and rhythm, have extremely low intercorrelations, no more than must be attributed to "g" alone. Spearman seems to attribute any apparent unitariness presented by musical ability together with most other special abilities, to past experience rather than to native aptitude.

Wing,¹⁰ using the modified factorial technique, which has been elaborated by Burt,¹¹ sought to examine the hypothesis of a general factor for musical ability, and to explore the influence of more specialized factors such as might obscure or interfere with the measurement of this general factor. He found that although success in music tests depends in part on general intelligence, nevertheless, there is a large group factor operating which appears to be specialized. The data analyzed were from seven of the more reliable recorded tests available on nine 10" gramophone records. These tests yielded the highest correlations with a preliminary

⁸ C. Spearman, *The Abilities of Man*, London: Macmillan (1927), p. 340.

⁹ *Ibid.*, p. 242.

¹⁰ H. D. Wing, "A Factorial Study of Musical Tests," *The British Journal of Psychology*, XXXI (1941), 342-344.

¹¹ C. Burt, *The Factors of the Mind*, London: University of London Press (1940), 494 pp.

battery of tests. They were standardized on over 2,000 boys and girls ranging in age from 8 to 18 from the London schools. The test problems were designed and graded in difficulty, the easiest being suitable for a child of eight, the hardest suitable to tax the capacities of a professional expert. The following tests were used: (1) detecting the number of notes played in a single chord, (2) detecting the change of a single note in a repeated chord, (3) detecting changes of notes in a short melodic phrase, (4) judging the more appropriate rhythmic accentuation in two versions of the same melody, (5) judging the more appropriate of two harmonized versions of the same melody, (6) judging the more appropriate mode of varying loudness (crescendo, decrescendo, sforzando, etc.) in two versions of the same melody, (7) judging the more appropriate phrasing (grouping of notes by pauses in two versions of a melody).

The results of Wing's¹² factor analysis disclosed three factors. The first factor is responsible for far more of the total variance than any other single factor. The assumption is that all seven tests may be treated as depending on a single common factor which is here designated "general musical ability." It is not a simple capacity, but is highly complex. Results imply that the author was testing one and the same group of mental processes. The second factor divides the seven tests into two sub-classes: the first includes all tests which requires the listener to judge the more appropriate musical arrangement, i.e., to decide the better phrasing, better rhythmical accentuation, better harmonization of a melody. The second sub-class comprises the three remaining tests in which the task is to perceive a change, i.e., to detect alterations in a melody, and in a chord, or in the number of tones or notes played simultaneously. The third factor showed appreciable saturations for two tests only; judging the better harmonization of a melody, and detecting the number of notes in chords. These were the only two tests which depended on listening to notes sounded simultaneously.

Drake,¹³ using Spearman's Tetrad-Difference Technique, in his

¹² *Ibid.*, 347-351.

¹³ R. W. Drake, "Factorial Analysis of Music Tests by the Spearman Tetrad-Difference Technique," *Journal of Musicology*, I (1939-1940), 6-16.

analysis of the Seashore tests, found that the criterion of division into two factors, one being a general factor and the other being a specific factor for each ability measured, but related to the general factor, is not satisfied. The purpose of the Seashore battery is to measure sensory attributes (though one is a test of memory) so that each test gives a measure of something quite different from any other. (The 1919 battery was used but the test of consonance was eliminated.) The results showed overlap of pitch with intensity, and overlap of time with intensity. Tonal memory and rhythm are independent except for saturation with the common factor. All the tests are related to one common factor. However, the fact that the tests have far more not common than common supports Seashore's "Theory of Specifics." On the other hand, the overlap shows that strictly isolated independent abilities are not measured.

Drake¹⁴ then used the following tests of his own: (1) retentivity, (2) musical memory, (3) pitch, (4) rhythm, (5) intensity, (6) time, (7) Seashore's Tonal Memory, and (8) Kwalwasser-Dykema Tonal Movement. He also administered a questionnaire and an intelligence test. He found more than one common factor among the eight tests. He said, however, that this was not final since all the tests cannot be accepted as valid tests of the talent. "We¹⁵ can only say that with the data at hand, two or more common factors are being measured by these tests." Even after partialling out age, intelligence, and training, he found group factors because the average deviation was still greater than its probable error. He found that musical memory should be given seven times the weight of pitch, and tonal memory six times the weight of pitch. Tests related only by the common factor to all eight are: musical memory, tonal memory, retentivity, intensity, and rhythm. There is one factor common to all eight tests and five additional factors which form group factors between tonal memory and tonal movement, pitch and intensity, pitch and tonal movement, musical memory and tonal memory, and intensity and time in order of magnitude. He found that there was a specific factor for pitch and considerable overlap between the two memory tests which suggests that pitch and at least one memory test be included in

¹⁴ *Ibid.*, p. 7.

¹⁵ *Ibid.*, p. 8.

any music testing program. He found that these two were distinct and fundamental in any music testing program. He maintains that this does not invalidate nor does it validate the multi-factor theory because it is impossible to determine whether such specific capacities should be summated in order to get a total estimate of musical talent or whether the common factor should be taken as the best measure of musicality.

Karlin¹⁶ used a comparable battery to that of Drake's but analyzed his data by means of Thurstone's method. His results show similarity in the factors emerging although these studies were made independently. He used 19 tests, 6 parts of Cattell's Intelligence test, 3 literary tests and a battery of 10 music tests. (Note: The writer states that he did not use Seashore's tests as they were too difficult for his subjects, but that he substituted 'similar tests.')

He found two oblique factors, a pitch or tonal sensitivity factor, and a memory factor. The memory factor may be analyzed as a gestalt for remembering musical passages, and a capacity for remembering isolated musical elements, a sensory or rote memory. He used Drake's tests of musical memory, and retentivity. He found that while the literary and intelligence tests correlated highly with each other and among themselves, of the 90 inter-correlations between the music tests and the rest of the battery, only 25 were as high as +0.10 and the mean correlation was +0.05. It appears that musical ability pertains largely to a field of its own, though there may be overlap between the more elemental components of intelligence and fundamental abilities peculiar to the music domain.

A few factorial studies, in the field of aesthetics, which deal chiefly with literary and artistic appreciation tests, include some music tests and are studied to see whether any or all of them may relate to a fundamental ability or 'appreciation phenomena.'

Eysenck¹⁷ found a general subjective factor of visual aesthetic appreciation (he used pictorial tests chiefly). This factor is not identical with "g" but rather it is a factor pertaining to a field of its own.

¹⁶ J. E. Karlin, "Music Ability," *Psychometrika*, VI (1941), 61-65.

¹⁷ H. J. Eysenck, "The General Factor in Aesthetic Judgments," *The British Journal of Psychology*, XXXI (1940), 94-102.

Dewar¹⁸ found that several different tests of visual aesthetic appreciation correlated together even after the influence of intelligence was eliminated. She correlated the Meier-Seashore Art Test The McAdory Art Test, tests for appreciation of applied art by Bulley, and Burt's Picture Post Card Test for artistic appreciation. She found a general factor which accounted for over 70 per cent of the variance and a second factor presenting bipolar types. Burt¹⁹ calls these types objective or classical, and subjective or romantic types. He has pointed out that much information can be gained not only by correlating tests but by correlating persons. In experiments in artistic appreciation, he suggests that the correlations of the individual tested with the performances of an ideal person taken as a standard, may provide a convenient measurement of that individual's capacities.²⁰ Inverted factor analysis yields more exact criteria for the determination of aesthetic types.

Williams, Winter, and Woods found the existence of a general factor which would be operative not only within one test, but between tests of different kinds of aesthetic appreciation (even after intelligence is eliminated). They found that capacity for literary appreciation correlates high with intelligence and to a lesser extent with pictorial and musical appreciation. Upon eliminating intelligence, significant partial correlations remain suggesting that aesthetic appreciation must also depend upon some general capacity of its own.

SUMMARY OF THE LITERATURE

A review of the literature on factorial analysis of musical ability presents diverging points of view. This cleavage may be due to the variety of kinds of tests employed and to the nature of the tests themselves.

Burt found evidence of a highly complex group factor for musi-

¹⁸ H. Dewar, "A Comparison of Tests of Artistic Appreciation," *British Journal of Educational Psychology*, VIII (1938), 29-49.

¹⁹ C. Burt, "The Psychology of Art," *How the Mind Works* (London: Allen & Unwin, 1933), 287-294.

²⁰ E. Williams, L. Winter, and J. M. Woods, "Tests of Literary Appreciation," *British Journal of Educational Psychology*, VIII (1938), 265-284.

cal ability, over and above the general factor for intelligence. He also found that persons as well as tests may be correlated and factorized by the inverted-factor technique.

Spearman doubts that there exists a group factor for music as well as other group factors and claims that experience rather than native ability accounts for unitariness in musical ability.

Wing found a large group factor which appears to be specialized.

Drake, in his analysis of the Seashore battery, found that the tests were related to one common factor and that some of the tests overlap. They do not measure isolated and independent abilities. In a further study, in which he used his own tests, he found one common factor and five additional factors.

Karlin found that musical ability pertains largely to a field of its own.

Eysenck found a general factor which he called aesthetic appreciation.

Dewar found a general factor which accounted for over 70 per cent of the variance, and a second factor presenting bipolar types.

Williams, Winter and Woods found the existence of a general factor which would be operative between tests of different kinds of aesthetic appreciation.

The question of whether the factor of general aesthetic appreciation which Eysenck calls the "t" factor is identical with Spearman's "g" factor or whether it relates exclusively to a field of its own has not been finally settled.

CHAPTER II

PROCEDURE OF EXPERIMENT

PART I

The music tests used in the factorial studies represent only a limited number of tests most of which are available in recorded form. In addition to these tests, a few of the other well known ones are: Schoen's²¹ battery of three tests (relative pitch, tonal sequence, and rhythm), Ortmann's²² battery of seven tests (pitch discrimination, pitch memory, time discrimination, fusion, rhythm memory, melodic memory, and harmonic memory), tests of music information and appreciation by Kwalwasser,²³ and tests of attitude and musical concepts by Hevner.²⁴ All of these tests were designed for the measurement of the reception of music, often called musicality, as contrasted with tests used in auditions for the reproduction of music, an aspect of music talent this study does not deal with.

It is important also to recognize the diversity of opinions and theories other modern authors hold regarding musicality. Vidor²⁵ stresses the importance of creative ability; thus she requests her subjects to fit melodies to rhythmic patterns dictated in taps and to complete melodic phrases. She holds that the relationship of

²¹ M. Schoen, "Tests of Musical Feeling and Musical Understanding," *Journal of Comparative Psychology*, V (1925), 31-52.

²² O. Ortmann, "Tests of Musical Talent," Peabody Conservatory of Music (unpublished), fide Schoen, *The Psychology of Music*, p. 280.

²³ J. Kwalwasser, *Test of Music Information and Appreciation*, Bureau of Education Research and Service, University of Iowa, Iowa City, Iowa (1927), 8 pp.

²⁴ K. Hevner, *Appreciation of Music and Tests for Appreciation of Music*, University of Oregon Publication, IV (1934), 73-150.

²⁵ Martha Vidor, *Was Ist Musikalität*, fide Mursell, *The Psychology of Music* (New York: W. W. Norton & Co., 1937), p. 302.