

14
A STUDY OF THE CONCURRENT VALIDITY OF THE
MUSICAL APTITUDE PROFILE

by

R. Leon Fosha

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Music in the Graduate
College of the State University of Iowa

3
August, 1964

9
Chairman:

Associate Professor Edwin Gordon

Co-Chairman:

Professor Leonard S. Feldt

ABSTRACT

The purpose of this study was to investigate the concurrent validity of the Musical Aptitude Profile, a new and unique musical aptitude test developed by Dr. Edwin Gordon at the State University of Iowa. The main problems of the study were 1) the investigation of the relationship between Musical Aptitude Profile test scores and a criterion of musical performance and 2) the investigation of the effects of formal music training on Musical Aptitude Profile test scores. The musical performance criterion consisted of student tape-recorded musical examples composed by the writer. Three auditors, working independently, evaluated each student performance utilizing a five point scale consisting of seven musical dimensions. A Pearson product-moment coefficient of correlation was computed between Musical Aptitude Profile test scores for each of nine groups consisting of band, string, and vocal students at the elementary, junior high, and senior high school levels. Obtained coefficients of correlation were corrected for homogeneity. Auditor reliabilities were estimated for any individual auditor and for the average rating from all three auditors.

Split-halves reliability coefficients for the complete test, based on the scores of 1418 combined musically

select and unselected students, were .943 for elementary, .953 for junior high, and .946 for senior high school students. Obtained validity coefficients were .399, .548, and .265 for bands, .251, .290, and .340 for strings, and .124, .266, and .343 for choirs for elementary, junior, and senior high groups respectively. Individual auditor reliability ranged from .617 to .909 while their average reliability ranged from .830 to .968 for the various performance groups.

The Musical Aptitude Profile was administered twice to 751 musically select and 647 unselected students in order to investigate the effects of formal music training on test scores. For the composite test score, musically select students at three school levels lost 3, gained 3, and lost 2 percentile rank points while unselected students at corresponding levels lost 3, lost 5, and lost 2 percentile rank points.

Test scores were correlated to determine rank order of musically select students from pre-test to post-test. Correlation coefficients were .829, .764, and .795 for elementary, junior high, and senior high school students respectively.

Because of the size of the study, factors such as the quality of training, student interest, and student

motivation could not be controlled. Despite these limitations, data obtained in each phase of the investigation can be considered contributory to the validity of the Musical Aptitude Profile as a test of musical aptitude.

Obtained reliability coefficients substantiate the battery as a reliable test for students enrolled in the upper elementary grades, junior high, and senior high school. Seven of the nine uncorrected validity coefficients, representing the relationship between test scores and the musical performance criterion, were statistically significant at the 5% level of confidence.

As coefficients cannot imply causation, it is not known whether high test scores resulted from musical aptitude or musical training. It seems logical, however, that high performance standards could not be achieved without the presence of comparatively high musical aptitude. Therefore, it is further concluded that the Musical Aptitude Profile, used cautiously, could be utilized for purposes of identifying musically talented students who will be successful in musical performance. The obtained evidence that formal music training has a negligible effect on Musical Aptitude Profile test scores tends to support this conclusion.

TABLE OF CONTENTS

	Page
I. Purpose of the Study.....	1
Introduction.....	2
Purpose of the Study.....	8
Problems of the Study.....	11
II. Previous Research.....	14
Introduction.....	15
The Wright Study.....	15
The More Study.....	18
The Brennan Study.....	22
The Stanton Study.....	25
The Taylor Study.....	26
The Wing Studies.....	30
Conclusions.....	33
III. Methods and Procedures.....	37
The Experimental Sample.....	38
The Testing Procedure.....	41
Development of the Criterion of Musical Achievement.....	42
Statistical Design and Analysis.....	48
IV. Presentation and Interpretation of the Data...	54
Introduction.....	55
Reliability of the Musical Aptitude Profile	56
Reliability of Auditor Evaluations.....	61
Coefficients of Concurrent Validity.....	61
A Comparison of Pre-Test and Post-Test Mean Score Differences for Musically Select and Unselected Students.....	67
Correlation of Pre-Test and Post-Test Scores for Musically Select Students.....	73
Test Intercorrelation Based on Pre-Test Data	76
V. Summary and Conclusions.....	82
Bibliography.....	91
Appendix.....	95

CHAPTER I

PURPOSE OF THE STUDY

I. INTRODUCTION

Interest in the development of tests to determine specific aptitudes enjoyed a substantial increase during the second decade of this century.¹ One of the earliest and best known attempts to measure aptitudes in music was that of Carl Seashore, whose Measures of Musical Talent² was published in 1919. Although musicians have remained skeptical, numerous other tests have appeared. Whybrew states:

Not only have several psychologists and teachers interested in this field brought forth such measures, but the number of tests presented by instrument manufacturers and distributors and by various types of institutions in the field of music should³ comprise a rather extensive bibliography.

The above author reserves his most poignant statement for his summation: "Needless to say, a wide range of merit

¹W. E. Whybrew, Measurement and Evaluation in Music, (Dubuque, Iowa: W. C. Brown and Company, 1962) p. 61.

²C. E. Seashore, Measures of Musical Talent, (New York: Columbia Phonograph Company, 1919).

³Whybrew, op. cit., p. vi.

is covered by these efforts."⁴

The content of an aptitude test reveals the nature of the constructs that support the author's hypothesis relative to the measurement of the specified trait.⁵ Hypotheses concerning the nature of musicality and its measurement have followed two divergent and fairly well defined lines.⁶

The first embodies the thought that since music consists of sound, man's ability to make critical judgments concerning the sound wave is indicative of musical talent. The group identified with this position has concerned itself primarily with human sensitivity to the various aspects of musical sound in terms of its physical characteristics such as frequency, intensity, time, and timbre.

The man who was perhaps most nearly representative of those who supported this hypothesis was Carl Seashore. He contended that, ". . . since music is conveyed from performer to listener by means of the sound wave, everything

⁴Ibid.

⁵A. Anastasi, Psychological Testing, 2nd edition, (New York: Macmillan Company, 1961) p. 137.

⁶Whybrew, op. cit., p. 86.

rendered and perceived as music can be expressed in terms of the physical attributes of the sound wave."⁷ It is important to stress that Seashore also recognized the existence of certain aspects of musical measurement other than those limited to the senses of pitch discrimination, loudness, duration, and timbre. The 1919 Seashore battery also included tests designed to measure tonal memory and consonance discrimination. The latter test was considered to be the least reliable sub-test of the battery and was subsequently dropped and replaced by a timbre test in the revised Measures of Musical Talents⁸ that appeared in 1939.

The validity of the Seashore tests is still a matter of question and remains a crucial issue in any assessment of their value. Seashore claimed validity for his individual tests in terms of their internal consistency.⁹ He contended that an outside criterion might not be as valid as the test.¹⁰

⁷Ibid., p. 87.

⁸C. E. Seashore, Measures of Musical Talents, (Camden, New Jersey: R.C.A. Manufacturing Company, 1939).

⁹C. E. Seashore, Psychology of Music, (New York: McGraw-Hill Book Company, Inc., 1938) p. 384.

¹⁰C. Leonhard and R. House, Foundations and Principles of Music Education, (New York: McGraw-Hill Book Company, Inc., 1959) pp. 343-344.

Although he encouraged others to investigate other types of validity for his tests, he never attempted directly to enhance his position beyond that of reaffirmation of his original tenets.

McLeish states:

The main question regarding the Seashore measures still remains: there is still the pressing need for a definitive study of validity of these revised measures and for a comparison of the different tests of musical ability using the same population subjects. . . . The Tests as now recorded have reached probably the peak of technical excellence so far as the presentation is concerned; it is rather shocking to find that so little is known about the validity of these tests after so many years of investigation.¹¹

In contrast to the "theory of specifics" supported by Seashore and his adherents, there are those who support an "omnibus" theory which follows the line of Gestalt Psychology.¹² The omnibus approach is based on the premise that musical aptitude cannot be evaluated within the limitations of the basic sensory capacities. Instead, they feel

¹¹J. McLeish, "Seashore Measures of Musical Talents, Revised Edition," Fourth Mental Measurements Yearbook, (Highland Park, New Jersey: Gryphon Press, 1953) 343.

¹²S. Bienstock, "A Review of Recent Studies of Musical Aptitude," Journal of Educational Psychology, 17:437, September, 1942.

that tests should be extended to include stimuli within a context more analogous to musical situations. Gestalt Psychologists believe that musical aptitude is dependent upon a general factor that is highly related to intelligent and aesthetic problem solving.¹³ In practice, this approach is most often supported by a contingent of Europeans who, using what is described as a "perceptual approach," find strong relationships between the two factors of intelligence and musical aptitude.¹⁴

Mursell and others who supported the omnibus theory contended that a measure of musical aptitude should be validated against some kind of musical activity.¹⁵ They felt that a test must function as a valid predictor of a musical behavior in order to be considered a valid measure of musical aptitude. In spite of this persistent view expressed by

¹³J. Mainwaring, "The Assessment of Musical Ability," British Journal of Educational Psychology, 17:96, June, 1947.

¹⁴S. Bienstock, Ibid., p. 431.

¹⁵Whybrew, Ibid., p. 90.

proponents of the omnibus theory, only a few limited concurrent and/or predictive validity investigations have been reported for the various published musical aptitude tests including the Seashore Measures of Musical Talent.

The question of aspects of validity remains the central issue in the controversy between the theory of specifics and the omnibus theory. Little progress has been made in distinguishing between the relative merits of the construct validity of these theories. Karlin states, however, "It seems reasonable to assume that the psychological attributes of the music domain will be compatible with the concept of measurement."¹⁶ Whybrew states:

In conclusion it must be said that no more than tentative assumptions can be made at present concerning the respective merits of these different approaches to the measurement of musical aptitude. No sufficient body of evidence is yet available to confirm or refute either theory. Further research is needed either to confirm or refute the claims of test authors for the validity of their measures.¹⁷

¹⁶J. E. Karlin, "Factor Analysis in the Field of Music," Journal of Musicology, 3:44, Summer, 1941.

¹⁷Whybrew, Ibid., p. 94.

II. PURPOSE OF THE STUDY

The impetus for this study was the appearance of a new test of musical aptitude developed at the State University of Iowa by Dr. Edwin Gordon. Entitled the Musical Aptitude Profile, the test is now in its fourth revision. The present form of the test represents a great many refinements resulting from analysis of data obtained in some 10,000 administrations in several midwestern cities. This study resulted from a need to investigate the various aspects of validity of this new battery.

The test author purports to deal with response to musical stimuli within a context of musical experience in the Musical Aptitude Profile. It is designed to measure, among other factors, aesthetic judgment. Although not directly concerned with fine discriminations of the sound wave at the basic sensory level, the test content is intended to naturally include this aspect of measurement.

The musical examples that comprise the battery are recorded on magnetic tape. Two stringed instruments, the violin and cello, are employed in the recording. The test is unique in respect to the use of these instruments as recording media. The examples themselves were composed by

the test author and do not represent extant music literature. The battery measures two types of responses, namely those of an objective (non-preference) and those of a subjective (preference) nature. The first four sub-tests measure the subject's ability to make absolute judgments based on the principle of musical variation. The last three sub-tests require subjects to state a preference for one song of a pair. A third, non-committal "in doubt" response is available for all seven sub-tests in order to ease tension, reduce guessing and, as a result, enhance reliability and validity aspects of the test battery.

The battery is described in further detail below:

Tonal Imagery

- Part 1. Melody: consists of paired musical examples, the second of which may or may not consist of a melodic variation of the first. Recorded by violin.
- Part 2. Harmony: similar in construction to the sub-test above. Here, however, the subject hears two voices instead of one but limits his judgment to a comparison of the lower voice in each of the two examples. Recorded by violin and cello.

Rhythm Imagery

- Part 1. Tempo: paired musical examples in which the endings may be identical or contrasted in tempo. Recorded by violin.

Part 2. Meter: paired musical examples that may be identical or contrasted in terms of altered meter or changes of rhythmic accent. Recorded by violin.

Musical Sensitivity

Part 1. Phrasing: paired examples, contrasted in terms of musical expression. Such contrast may consist in changes of tone, tempo, dynamics, or any combination of these variables. Recorded by violin and cello.

Part 2. Balance: paired examples in which the endings are contrasted in terms of rhythmic and melodic content. Recorded by violin.

Part 3. Style: paired examples, the tempo of the first being contrasted with that of the second, thus effecting a contrast of musical character. Recorded by violin.

The present form of the test is unique in construct validity and practicality. The items supposedly require practically no formal musical achievement on the part of the subject. It is the writer's opinion that, at most, an unsophisticated exposure to music should be sufficient for understanding the questions but not necessary for choosing the correct answer. It is the only existing test that measures responses to melody, harmony, rhythm, and musical sensitivity in such a musically indigenous fashion.

III. PROBLEMS OF THE STUDY

This study was principally concerned with an investigation of the concurrent validity of the Musical Aptitude Profile. Related problems concerned with the influence of musical training on scores and test reliability were also investigated. In addition, the relationship among the various sub-tests was investigated through intercorrelation procedures. The main problems of the study were:

1. The investigation of the relationship of test scores and a criterion of musical achievement as defined by the writer.
2. The investigation of the stability of the battery in terms of the influence of musical training on test scores.

Secondary problems of the study bearing on the Musical Aptitude Profile were:

1. The investigation of the reliability of the test battery.
2. The intercorrelation among the various tests.

IV. DEFINITION OF TERMS

Concurrent validity. The extent to which test scores