AN INVESTIGATION OF THE VALIDITY OF THE GORDON MUSICAL APTITUDE PROFILE

by

Vernon V. Tarrell

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

August, 1964

Co-Chairmen: Associate Professor Edwin Gordon
Associate Professor Leonard S. Feldt

The problem of the study was to determine the relationship between Gordon Musical Aptitude Profile test scores and musical performance as defined by a special criterion of success in instrumental or vocal music. The study also investigated the reliability of the test battery.

Nine hundred students were selected from 1487 who were administered the Musical Aptitude Profile in order to study the relationship between test scores and defined musical success. The students were grouped into three grade classifications: grades 4 through 6, grades 7 through 9, and grades 10 through 12. They were then placed into one of three categories: 1) Select students in instrumental music. This category consisted of 100 brass and woodwind players from each of the three grade classifications, comprising a total of 300 students; 2) Select students in vocal music. This category consisted of 100 vocal students from each of the three grade classifications, and totalled 300 students; 3) Unselected students. This category was composed of 100 students from each of the three grade classifications, and totalled 300 students.

Musical examples, specially composed by the investigator for each grade classification and medium of performance, were distributed to the music supervisors in January, 1963. Student performances of the musical examples were tape recorded

during April and May, 1963. A committee of three professional musicians then made independent evaluations of the performances, employing the following criterion of success in music: 1) suitability of tone quality; 2) suitability of intonation; 3) suitability of meter; 4) suitability of tempo; 5) suitability of phrasing; 6) suitability of rhythm; and 7) suitability of style. A five-point scale was used to rate each characteristic of the student's performance; "Poor," "Below Average," "Average," "Above Average," and "Excellent."

The ratings by each judge were summed and an average score was assigned to each student's performance. Performance scores were correlated with each of the three part scores and with the composite test score of the Musical Aptitude Profile to determine validity coefficients.

Reliability coefficients for all test scores in the battery were computed for each grade and corrected through the use of the Spearman-Brown formula. Also, intercorrelation coefficients for all test scores were determined.

Within the limitations of this investigation and under present standards, the following tentative conclusions seem warranted:

1) The Musical Aptitude Profile will materially assist teachers in the identification of students enrolled in grades 4 through 12 who might profit most from special music instruction.

2) The Musical Aptitude Profile appears to be a consistently stable test battery for all students who are enrolled in grades 4 through 12.

3) Intercorrelation analyses of the various sub-tests in the battery reveal substantial relationships for each with the composite test. Relationships between sub-tests within each part test are also relatively high; however, there was no evidence that any particular sub-test was sufficiently related to any other sub-test as to duplicate it in function.

Abstract Approved: _____, Dissertation Supervisor

Title and Department: _____

Abstract Approved: _____, Dissertation Co-Chairman

Title and Department: _____

Date: _____

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

CHAPTER I

I. INTRODUCTION

In 1919 Carl E. Seashore constructed a battery of tests in which he proposed to measure aptitude in music through the sound wave and its four physical attributes: frequency, amplitude, duration, and form.[1] The battery consisted of tests of Pitch, Time, Intensity, Tonal Memory, and Consonance. Six years later a test of Rhythm was added. Seashore, Saetveit, and Lewis published a revised version of the test in 1939.[2] The Consonance test was dropped from the battery because of its low reliability, and was replaced with a Timbre test.

Seashore held his tests to be valid measures of musical talent. Lundin, however, reported a wide range of concurrent validity coefficients* in a summary of several independent

---

[1]Carl E. Seashore, "Seashore Measures of Musical Talent," New York: Columbia Phonograph Co., 1919.

[2]J. G. Saetveit, D. Lewis, and Carl E. Seashore, "Revision of the Seashore Measures of Musical Talents," University of Iowa Studies on Aims and Progress of Research, 65, Whole 388 (1940).

---

*A test is said to have concurrent validity to the extent that the test scores correlate positively with other measures of the same trait or ability, the latter being obtained about the same time as the test scores.

investigations of the test battery.[3] Seashore reported what he called internal validity coefficients for his tests; that is, it was demonstrated that all items in the Pitch test, for instance, were highly intercorrelated. The same might be said for the other tests.

Heinlein attacked the so-called "atomistic" approach to musical aptitude testing. In a study of the Seashore Consonance test,[4] he criticized it for its lack of construct validity,* stating that the paired interval comparison was an inadequate method of testing the sense of consonance. In another study,[5] Heinlein attacked the wisdom of trying to diagnose musical talent through the use of an "identification" type method such as Seashore used in his Tonal Memory test. Seashore's test

---

[3]Robert W. Lundin, "A Preliminary Report on Some New Tests of Musical Ability," Journal of Applied Psychology, No. 28 (1944), pp. 393-6.

[4]C. F. Heinlein, "An Experimental Study of the Seashore Consonance Test," Journal of Experimental Psychology, 8 (1925), pp. 408-33.

[5]C. F. Heinlein, "A Brief Discussion of the Nature and Function of Melodic Configuration in Tonal Memory with Critical Reference to the Seashore Tonal Memory Test," Pedagogical Seminar, Journal of Genetic Psychology, 35 (1928), pp. 45-61.

---

*A test may be said to have construct validity if the tasks included in the test constitute an acceptable operational definition of the author's basic concept of the abilities to be measured.

required the subject to listen to a pattern of from three to five notes. The subject was then asked to determine which note had been altered when the pattern was replayed.

Both Heinlein and Mursell believed that the Measures of Musical Talent were really sensory discrimination tests. In effect, they argued that Seashore's definition of musical aptitude failed to include elements which they regarded as crucial. They felt that musical aptitude tests should include actual musical sounds in the test items. This "omnibus" theory[6] was in direct contrast to the Seashore "atomistic" approach.[7] Stimulated by these opposing views, researchers have conducted a wide variety of studies in the nature and development of musical aptitude. The primary goal of much of this research has been the early identification of those showing unusually high levels of musical aptitude.

The development of a valid instrument for measuring aptitude in music would, without doubt, have far-reaching, positive implications for the improvement of music education in the nation's schools. The teacher of general music could provide more effective teaching and evaluation of student growth in music

---

[6]James L. Mursell, Psychology of Music (New York: W. W. Norton and Co., Inc., 1937).

[7]Carl E. Seashore, Psychology of Music (New York: McGraw-Hill Book Co., 1938).

if aptitude could be identified at an early age. Music educators who rely upon selectivity as a means of determining participation in the vocal and instrumental music programs of their schools would find an objective test of musical aptitude a valuable aid in such selection. For the student, such a test could well be the best opportunity for him to make known to himself and others his aptitude for music, an aptitude that might otherwise go unnoticed.

II. PURPOSE OF THE STUDY

The recent appearance of a new test of musical aptitude, the Musical Aptitude Profile by Dr. Edwin Gordon of the State University of Iowa,[8] prompted the present study. An examination of the test battery revealed that because of several important new approaches to testing musical aptitude it merited further investigation. The unique aspects of this test are as follows: 1) The tests contain original musical examples. When musical materials have been used, other test authors have used familiar examples written by recognized composers. 2) The test items were tape recorded by professional musicians of international reputation. No other test author

---

[8] Edwin Gordon, Musical Aptitude Profile, Boston: Houghton Mifflin Co., 1965.

has utilized professional musicians for this purpose. 3) The violin and cello were used as recording instruments. String instruments as musical stimuli have heretofore not been employed in connection with tests of musical aptitude. 4) The test items demand a variety of discriminations and judgments. The specific techniques that are employed are found in no other standardized test of musical aptitude. 5) The Musical Aptitude Profile is the only American standardized test to include preference tests in its battery. Wing, in England, has published a test battery which includes preference tests.[9]

The present battery of tests consists of seven subtests. It is the fifth revision of the original battery. The test is the culmination of many refinements that resulted from data obtained through more than eight thousand test administrations in Racine and Wausau, Wisconsin; Centerville, Shenandoah, and Maquoketa, Iowa; and Sandusky, Ohio.[10]

Since the results suggest that the Gordon battery holds considerable promise for the early identification of the highly talented individual, it was desirable to examine the battery in more detail to gather additional data that would bear on the

---

[9]H. D. Wing, "A Revision of the Wing Musical Aptitude Test, Journal of Research in Music Education, 10, No. 1 (1962), pp. 39-46.

[10]General Manual, Musical Aptitude Profile, Boston: Houghton Mifflin, 1965.

validity of the test. In order to accomplish this, one specific aspect of validity was investigated: namely, concurrent validity.

III. STATEMENT OF THE PROBLEM

The problem of this study was to determine the relationship between obtained test scores of the Musical Aptitude Profile and quality of a tape recorded musical performance. Specifically, the problem of the study was:

> To determine the relationship between Musical Aptitude Profile test scores and qualitative measures of musical performance, as defined by a special criterion of success in instrumental or vocal music.

The study also investigated the following issues: 1) The reliability of the test, by obtaining coefficients of internal consistency. 2) The intercorrelation between the various tests comprising the Musical Aptitude Profile.

IV. LIMITATIONS OF THE STUDY

The conclusions drawn from this study were limited by the following factors over which complete control could not be exercised:

1) No contact with the students prior to the administration of the test battery was possible. Motivation and interest for taking the Musical Aptitude Profile was developed largely

through the efforts of school officials and music supervisors.

2) In some instances crowded seating arrangements were unavoidable. It is possible that these conditions, where they existed, hampered the students' efforts to achieve their best scores on the test battery.

3) No contact with the students prior to the tape recording of the musical performances was possible. Personal motivation by the investigator was therefore limited.

4) It was not feasible to personally supervise the preparation of the musical material for the tape recorded performances. Since differences in the amount and quality of preparation inevitably resulted, it was impossible to determine whether a truly representative sample of each student's musical behavior was obtained.

5) Although school officials generally provided excellent facilities for tape recording student performances, none of the rooms was soundproof. Consequently, distractions resulting from noises originating outside the recording rooms were frequent, and resulted in less than ideal recording conditions.

6) Judgment of each student's musical behavior was limited by how well or poorly he played or sang the musical material for the one tape recording.

7) Evaluation of each student's musical behavior was limited to the extent of subjective agreement among the

individual judges.

V. THE MUSICAL APTITUDE PROFILE

The Musical Aptitude Profile consists of the following parts: Part I, Tonal Imagery; Part II, Rhythm Imagery; and Part III, Musical Sensitivity. In Part I there are two sub-tests, the first of these being Melodic Variations. This consists of two musical examples for each test item, the second of which may or may not involve a melodic variation of the first. The response required of the subject is non-preferential in nature; that is, the subject is not asked to indicate the two examples he prefers, but instead whether or not a change in the melody has occurred in the repetition of the example. There is further provision for an "in doubt" response if the subject is not sure of his judgment of a particular item. The violin is the medium of performance, and this sub-test has a total of twenty pairs of items. The second sub-test of Tonal Imagery is a test of Harmonic Variations, and is similar in construction to that of the first sub-test. It includes twenty pairs of items which are tape recorded by a violinist and a cellist. The subject is asked to compare the part for the lower instrument only, since the upper part is played exactly the same both times.

Part II, that of Rhythmic Imagery, also includes two sub-tests, each consisting of twenty pairs of items which are tape recorded by a violinist. The first sub-test, Tempo, is composed of paired musical examples, in which the endings may or may not involve a change of tempo in its repetition. The second test, Meter, consists of paired musical examples which may, upon repetition, be either identical or involve changes of meter.

Part III of the Gordon battery includes three sub-tests of Musical Sensitivity. These sub-tests are preferential in nature; that is, they call for expressions of personal preference on the part of the subject. The scoring of these items was arrived at by asking ten professional musicians to take the items. Unless nine of the ten professionals agreed upon a given preference, the item was discarded. Further study of the items was made through the techniques of analyses after each administration of the test battery.

The first sub-test in Part III is a measure of Phrasing. The musical examples are tape recorded by a violinist and a cellist, and present contrasts in musical expression. The second sub-test is that of Balance, in which the musical examples played by a violinist have contrasting rhythmic and melodic endings. The third sub-test is a measure of the subject's ability to interpret Style. The musical examples are

recorded by a violinist, and reflect differences in style of performance through changes of tempo.

The directions for marking the answers are recorded on tape along with the test items. They are readily understood by a student who has been exposed to music, but who has not necessarily had training in music. The test battery has been so constructed that each of the three parts may be easily and completely administered during a fifty-minute class period.

Eleven test scores are obtained from the student. For convenience, each test has been given an abbreviated title. The test of Tonal Imagery may be identified by the letter T; similarly, R represents Rhythm Imagery and S represents Musical Sensitivity. The letter C indicates the composite test scores for the complete test battery. The number subscript indicates the particular sub-test within the part test. For example, T_1 represents the test of Melodic Variations in Part I, Tonal Imagery.