A PILOT STUDY TO INVESTIGATE CONSTRUCTS OF MUSICAL APTITUDE

by

Edward L. Rainbow

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

February 1963

Co-Chairman: Assistant Professor Edwin Gordon

Co-Chairman: Associate Professor Leonard Feldt

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

CHAPTER I

PURPOSE OF THE STUDY

I. INTRODUCTION

In the second decade of the twentieth century, Dr. C. E. Seashore envisioned the development of musical aptitude tests that would:

> ...be used as as (sic.) group tests, for the purpose of a rough preliminary sifting in the school room. This will reveal the unusually bad as well as the unusually good; and both of these classes deserve individual treatment. Such tests may eliminate the helplessly unmusical and save them from an intolerable imposition of musical requirements; but their real value is in finding the gold in the dross. One gifted child found early, investigated, and encouraged is a great reward.[1]

In 1919, Dr. Seashore published the first standardized musical aptitude test.[2] He asserted that this test was the psychological tool that would enable man to find "the gold in the dross."

In the following four decades additional tests designed to measure musical aptitude were developed and

---

[1] C.E. Seashore, Measures of Musical Talent. (New York:G. Schirmer, 1915), p. 19-20.

[2] _____. Measures of Musical Talent. (New York:Columbia Phonograph Co., 1919).

made available to music educators and psychologists. The most noted of these tests were the Kwalwasser-Dykema Music Tests,[3] The Tilson-Gretsch Musical Aptitude Test,[4] the Drake Musical Aptitude Tests[5], the Gaston Tests of Musicality[6], the Wing Standardized Tests of Musical Intelligence[7], and the 1939 revision of the Seashore Measures of Musical Talents.[8]

Music educators and psychologists generally agree that a need exists for instruments that will efficiently measure musical aptitude. Both of these groups, however, have been critical of the current status of musical

---

[3]J. Kwalwasser and P.W. Dykema, Kwalwasser-Dykema Music Tests. (New York:C. Fischer, 1930).

[4]L.M. Tilson, The Tilson-Gretsch Musical Aptitude Test, (Chicago:Public School Music Department, Fred Gretsch Mfg. Co.) 1941.

[5]R.M. Drake, Drake Musical Aptitude Tests. (Chicago:Science Research Associates, 1954).

[6]E. Thayer Gaston, Test of Musicality 4th edition. (Lawrence, Kansas:O'Dell's Instrument Service, 1958).

[7]H. Wing, Wing Standardized Tests of Musical Intelligence, (London:National Foundation for Educational Research in England and Wales).

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

aptitude testing.

The following statements are, in general, representative of the criticisms leveled at musical aptitude tests.

Charles Leonhard, a music educator, has stated:

> The decision to use or not to use tests of musical capacity requires mature professional judgment and should be made only after consideration of all the factors involved. In view of the extreme complexity of the problem, the conclusion of this writer is that the present state of development of tests of musical capacity does not warrant their use as the sole determinant in screening students for specialized musical opportunities.[9]

The board of directors of the federated state units of the Music Educators National Conference has issued the following statement relating to musical aptitude testing:

> Because of our inability to communicate clearly the values of music, our lack of instruments to measure musical aptitude precisely, and our lack of agreement on what constitutes a sequential, developmental music curriculum, counselors cannot counsel as wisely in music as they do in other academic subjects.[10]

---

[9] C. Leonhard, "Evaluation in Music Education", The Fifty-Seventh Yearbook of the National Society for the Study of Education, Part I. Chicago: University of Chicago Press, 1958, p. 322.

[10] "What are Music Educators Thinking and Saying: a Digest of the Complete report of the Interim Meeting of the Board of Directors of the Federated State Units of the MENC." Music Educators Journal. 48:2 Nov. - Dec.. 1961, p. 60.

Eminent authorities in the field of psychological measurement have voiced criticism of musical aptitude tests. Rothney has stated:

> The best results obtained so far are indicated by such small coefficients of correlation between test scores and criteria that prediction of an individual's later performance in the area christened by the author is little better than chance.[11]

Anastasi has offered the following observations:

> The development of tests specifically designed for measuring aesthetic abilities... has been slow and sporadic. Little progress in the testing of artistic, musical or literary aptitudes has been made since early 1940's. In number, scope, and technical refinements, tests in this area have lagged far behind other aptitude tests.[12]

Much of the criticism leveled at musical aptitude tests by music educators and psychologists has arisen because of their refusal to accept oft-used criteria of validity. Thus reservations regarding the available evidence can probably be attributed, at least in part, to the lack of a generally acceptable criterion of musical

---

[11] J.W. Rothney, P.J. Danielson, and R.A. Heiman, Measurement for Guidance (New York: Harper and Bros., 1959), p. 330-331.

[12] A. Anastasi, Psychological Testing, 2nd edition (New York: Macmillan Co., 1961), p. 400.

aptitude itself.

The content of an aptitude test reveals the hypothesis upon which that test was developed. An examination of the labels of the subtests of several selected musical aptitude tests reveals striking similarities in these hypotheses among the various test constructors.

The Seashore test, in its original form, had five subtests which were purported to measure the Sense of Pitch Discrimination, the Sense of Intensity, the Sense of Time, the Sense of Consonance, and the Sense of Tonal Memory. In the 1939 revision of this battery, the Sense of Consonance test was deleted, the Sense of Intensity test was renamed the Sense of Loudness, and tests to measure sensitivity to rhythm and timbre were added. In its revised state, this battery then contains tests to measure the Sense of Pitch Discrimination, the Sense of Loudness, the Sense of Rhythm, the Sense of Time, the Sense of Timbre, and the Sense of Tonal Memory.

The Kwalwasser-Dykema Test contains ten subtests which the authors propose as measures of senses of Pitch, Intensity, Time, Rhythm, Timbre, Tonal Memory, Tonal Movement, Melodic Taste, Pitch Imagery, and Rhythmic Imagery. The Drake test has two subtests that are

intended to measure Musical Memory and Rhythm while the Tilson-Gretsch test presents tests concerning Pitch, Intensity and Time Sensitivity, and Tonal Memory. The Wing battery contains seven subtests; Chord Analysis, Pitch Change, Memory, Rhythmic Accent, Harmony, Intensity, and Phrasing. In general, constructors of musical aptitude tests seem to agree that musical aptitude can be assessed by measuring such functions as: the sense of pitch, the sense of tonal memory, the sense of rhythm, the sense of time, and the sense of intensity.

Anastasi has stated:

> ...the content of aptitude and personality tests can do little more than reveal the hypothesis that led the test constructor to choose a certain type of content for measuring a specified trait. Such hypotheses need to be empirically confirmed to establish the validity of the test.[13]

In the case of the above mentioned musical aptitude tests, such hypotheses have not been empirically confirmed. Perhaps this is the underlying reason why the development of musical aptitude tests has lagged behind other aptitude tests.

Perhaps, as Rothney states:

---

[13] Ibid., p. 137

> The time has come for test builders to reexamine their basic premeses and techniques. Continuation of the usual timeworn processes of test construction that have proved to be almost sterile seems not to be justified....[14]

The writer concurs with Rothney and believes that such a reexamination is necessary. This reexamination should center on an empirical investigation of the constructs of musical aptitude. The writer believes that the following procedures will, in fact, give the constructors of future musical aptitude tests direction for the establishment of construct validity. The following sequence is suggested: (1) an attempt should be made to develop a definition of musical aptitude against which subjects could be evaluated, (2) a list of variables hypothesized to be associated with musical aptitude should be developed, (3) subjects possessing high and low levels of this defined aptitude should be identified by judges, (4) analyses should be made to discover the strength of the relationship between levels of defined musical aptitude and each of the hypothesized variables.

---

[14] Rothney, op. cit., p. 331.

II. PURPOSE OF THE STUDY

The purpose of this study was to give the constructors of future musical aptitude tests direction for the establishment of construct validity. The steps enumerated above were applied. By using these procedures it was possible to investigate empirically the relationship of each of the following variables to level of musical aptitude:

1. Pitch Discrimination
2. Tonal Memory
3. Rhythm
4. Musical Memory
5. Academic Intelligence
6. School Achievement
7. Sex
8. Chronological Age
9. Musical Achievement
10. Musical Training
11. Home Enrichment
12. Interest in Music
13. Participation in Music by Relatives
14. Socio-Economic Background

If an identification of the variables which con-

tribute to musical aptitude can be made, it may then be possible to try to develop tests which will efficiently measure these variables.

III. PROBLEMS OF THE STUDY

The primary problems of this study were as follows:

(1) To estimate the mean differences on each of the fourteen variables mentioned above among groups of students who possessed differential levels of defined musical aptitude.

(2) To determine the degree of interrelationship of the fourteen variables and the additional variable, defined musical aptitude.

(3) To determine which variables, as a group, most clearly reflected the factors which constitute defined musical aptitude.

The data appropriate to each of these problems was separately accumulated for three different grade levels: elementary school (grades 4 to 6), junior high school (grades 7 to 8), and high school (grades 9 to 12).

IV. DEFINITION OF TERMS

The American Psychological Association recognizes four aspects of validity. These aspects are content validity, predictive validity, concurrent validity, and construct validity.[15] For the purpose of this study the four aspects of validity will be defined as follows:

Content validity. How well the content of the test samples the subject matter about which the conclusions are to be drawn.[16]

Predictive validity. How well predictions made from the results of testing are confirmed by evidence gathered at a subsequent time.[17]

Concurrent validity. How well test scores correspond to measures of parallel criterion performance or status.[18]

---

[15]Joint committee of American Psychological Association, American Educational Research Association, and National Council on Measurements used in Education. Technical Recommendations for Psychological Tests and Diagnostic Techniques. Washington:American Psychological Association, 1954, p. 13.

[16]Ibid., p. 13

[17]Ibid., p. 13

[18]Ibid., p. 14