

A VALIDITY INVESTIGATION OF THE
DRAKE MUSICAL APTITUDE TESTS

by

John Woodring Ferrell

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 1961

Chairman: Professor Edwin E. Gordon
Co-Chairman: Professor Leonard Feldt

ACKNOWLEDGMENTS

The writer wishes to express his indebtedness and appreciation to Dr. Edwin E. Gordon for his constructive criticism and guidance which enabled the completion of this investigation.

To Dr. Neal Glem for his interest and encouragement, the writer is deeply grateful.

The writer also wishes to acknowledge Dr. Leonard Feldt for his assistance with the statistical portion of this study.

In addition, gratitude is extended to the administrators, faculties and students of Charles City, Shenandoah, and Washington, Iowa, who participated in this investigation. Without their cooperation this study would not have been possible.

Above all the writer wishes to express appreciation to his wife for her understanding and assistance through the course of this study.

TABLE OF CONTENTS

CHAPTER	PAGE
I. THE PROBLEM: STATEMENT AND ANALYSIS	1
Introduction and Purpose	1
Musical Aptitude Tests: Criteria of	
Validity	3
Summary	12
Discussion	14
Statement of the Problem	16
Limitations of the Study	19
II. A REVIEW OF THE RELATED STUDIES	21
Introduction	21
The Stanton Study	22
The William Larson Study	25
The Lawshe and Wood Study	26
The Wing Study	32
Criticisms and Implications of the Studies .	33
III. METHODS AND PROCEDURES	36
Measures of Musical Aptitude: Description .	36
Selection of the Experimental Subjects . . .	37
Definition of the Groups	38

CHAPTER	PAGE
Test of Musical Concepts	40
Selection of Music-Participant Subjects .	41
Selection of Non-Music-Participant Subjects	43
Hypotheses	44
Administration of the Tests	45
Statistical Design and Analysis	47
Analysis of Variance	47
Comparisons of groups with similar understanding of musical concepts . .	48
Comparisons of groups with similar musical aptitude	49
Mixed comparisons	49
Comparisons of educational development with musical aptitude test scores . .	50
Analysis of Covariance	51
IV. DATA AND RESULTS OF THE STUDY	52
Within-School-Group and Total-Group Mean Scores	52
Analysis of Variance	56
Comparisons of Groups with Similar Understanding of Musical Concepts . . .	56

CHAPTER	PAGE
Comparison of Groups with Similar	
Musical Aptitude	63.
Mixed Comparisons	69
Comparison of Educational Development	
and Musical Aptitude	69
Analysis of Covariance	76
Comparison of Groups with Similar	
Understanding of Musical Concepts . . .	76
Comparison of Groups with Similar	
Musical Aptitude	85
Mixed Comparisons	96
V. SUMMARY AND CONCLUSIONS	104
Experimental Procedures	104
Observations and Conclusions	106
BIBLIOGRAPHY	116
APPENDIX A	127
APPENDIX B	135

LIST OF TABLES

TABLE	PAGE
I. William Larson Study: Average Percentile Ranks of Subjects on the Six Subtests of the Seashore Measures of Musical Talent	27
II. Lawshe and Wood Study: Number of Students Taking Each Test in Total Group and in Musical Organizations	29
III. Lawshe and Wood Study: Number, Median Scores and Probable Error Choral Group vs. Non-Music-Participant Group	30
IV. Lawshe and Wood Study: Number, Median Scores and Probable Error Instrumental Group vs. Non-Music-Participant Group	31
V. Within-School-Group and Total-Group Mean Scores: Groups A, B, C, D, and E from Schools I, II, and III on the Test of Musical Concepts	53
VI. Within-School-Group and Total-Group Mean Scores: Groups A, B, C, D, and E from Schools I, II, and III on the Drake Musical Memory Test	54

TABLE	PAGE
VII. Within-School-Group and Total-Group Mean Scores: Groups A, B, C, D, and E from Schools I, II, and III on the Special Test of Musical Memory	55
VIII. Within-School-Group and Total-Group Mean Scores: Groups A, B, C, D, and E from Schools I, II, and III on the Iowa Tests of Educational Development	57
IX. Summary Table of Analysis of Variance for the Drake Musical Memory Test: Groups A, B, and D	58
X. Comparison of the Difference of Mean Scores of Experimental Groups A, B, and D for the Drake Musical Memory-Test	59
XI. Summary Table of Analysis of Variance for the Drake Musical Memory Test: Groups C and E	60
XII. Summary Table of Analysis of Variance for the Special Test of Musical Memory: Groups A, B, and D	61

TABLE	PAGE
XIII. Comparison of the Difference of Mean Scores of Experimental Groups A, B, and D for the Special Test of Musical Memory	62
XIV. Summary Table of Analysis of Variance for the Special Test of Musical Memory: Groups C and E	64
XV. Summary Table of Analysis of Variance for the Drake Musical Memory Test: Groups B and C	65
XVI. Summary Table of Analysis of Variance for the Drake Musical Memory Test: Groups D and E	66
XVII. Summary Table of Analysis of Variance for the Special Test of Musical Memory: Groups B and C	67
XVIII. Summary Table of Analysis of Variance for the Special Test of Musical Memory: Groups D and E	68
XIX. Summary Table of Analysis of Variance for the Drake Musical Memory Test: Groups C and D	70

TABLE	PAGE
XX. Summary Table of Analysis of Variance for the Special Test of Musical Memory: Groups C and D	71
XXI. Summary Table of Analysis of Variance for the Iowa Tests of Educational Development: Groups A and B	72
XXII. Summary Table of Analysis of Variance for the Iowa Tests of Educational Development: Groups A and B	73
XXIII. Summary Table of Analysis of Variance for the Iowa Tests of Educational Development: Groups B and D	74
XXIV. Summary Table of Analysis of Variance for the Iowa Tests of Educational Development: Groups C and E	75
XXV. Summary Table of Analysis of Variance for the Iowa Tests of Educational Development: Groups B and C	77
XXVI. Summary Table of Analysis of Variance for the Iowa Tests of Educational Development: Groups D and E	78

TABLE	PAGE
XXXVI. Adjusted Means for Drake Musical Memory Test: Groups B and C	90
XXXVII. Summary Table of the Analysis of Covariance for the Drake Musical Memory Test: Groups B and C	91
XXXVIII. Adjusted Means for Drake Musical Memory Test: Groups D and E	92
XXXIX. Summary Table of the Analysis of Covariance for the Drake Musical Memory Test: Groups D and E	93
XL. Adjusted Means for Special Test of Musical Memory: Groups B and C	94
XLI. Summary Table of the Analysis of Covariance for the Special Test of Musical Memory: Groups B and C	95
XLII. Adjusted Means for Special Test of Musical Memory: Groups D and E	97
XLIII. Summary Table of the Analysis of Covariance for the Special Test of Musical Memory: Groups D and E	98
XLIV. Adjusted Means for Drake Musical Memory Test: Groups C and D	99

TABLE	PAGE
XLV. Summary Table of the Analysis of Covariance for the Drake Musical Memory Test: Groups C and D	100
XLVI. Adjusted Means for Special Test of Musical Memory: Groups C and D	101
XLVII. Summary Table of the Analysis of Covariance for the Special Test of Musical Memory: Groups C and D	102
XLVIII. Raw Scores Received by Subjects in Experimental Group A, Schools I, II, and III, on the Test of Musical Concepts, the Drake Musical Memory Test, and the Raw Composite Scores on the Iowa Tests of Educational Develop- ment	136
IL. Raw Scores Received by Subjects in Experimental Group B, Schools I, II, and III, on the Test of Musical Con- cepts, the Drake Musical Memory Test, and the Raw Composite Scores on the Iowa Tests of Educational Development	138

TABLE

PAGE

I. Raw Scores Received by Subjects in
 Experimental Group C, Schools I, II,
 and III, on the Test of Musical Con-
 cepts, the Drake Musical Memory Test,
 and the Raw Composite Scores on the
 Iowa Tests of Educational Development . . . 140

LI. Raw Scores Received by Subjects in
 Experimental Group D, Schools I, II,
 and III, on the Test of Musical Con-
 cepts, the Drake Musical Memory Test,
 and the Raw Composite Scores on the
 Iowa Tests of Educational Development . . . 142

LII. Raw Scores Received by Subjects in
 Experimental Group E, Schools I, II,
 and III, on the Test of Musical Con-
 cepts, the Drake Musical Memory Test,
 and the Raw Composite Scores on the
 Iowa Tests of Educational Development . . . 144

LIST OF FIGURES

FIGURE	PAGE
1. Significant Differences and Interactions Obtained through the Analysis of Variance of the Means of the Groups for the Drake Musical Memory Test, the Special Test of Musical Memory, and the Iowa Tests of Educational Development	80
2. Significant Differences and Interaction Obtained through the Analysis of Covari- ance of the Means of the Graphs for the Drake Musical Memory Test and the Special Test of Musical Memory	103

CHAPTER I

THE PROBLEM: STATEMENT AND ANALYSIS

I. INTRODUCTION AND PURPOSE

The validation of musical aptitude tests became a major concern of music educators and psychologists following the publication of the Seashore Measures of Musical Talent in 1919. Since that time, additional musical aptitude tests have been developed and numerous investigations have been undertaken to determine their validity. Many music educators, however, are not entirely satisfied with the techniques used to determine the validity of these tests and are unwilling to fully accept the test results. In fact, much of the hesitancy of music educators in utilizing music aptitude test results can be attributed to a distrust of the methods of test validation.

The selection of a criterion representing a satisfactory measurement of musical aptitude constitutes a principal problem and has caused considerable controversy among music educators. This debate is partly founded on the subjectivity of various criteria used in validity investigations. Validation studies which have been designed to use teacher estimates of musical characteristics as a

criterion may be contaminated by extraneous factors to an undetermined extent. For example, teachers' estimates of levels of ability in applied music or sight singing are only as consistent as the teachers' standards for evaluation. Similarly, teachers' judgments of musical aptitude are analogous only if the teachers agree on the meaning of the term "musical aptitude." What is reflected in a validity coefficient based upon teacher estimates of the subjects' capabilities depends upon the individual teacher's understanding and interpretation of the criterion. As Drake has pointed out, the cogency of teacher estimates relies upon

... the extent to which the rater or judge is able accurately to evaluate and rate the individuals taking the tests, and also how well the rater can define true ability apart from length of musical training, general intelligence, personality and effort¹

The purpose of this study was to investigate and clarify the power of a music aptitude test to discriminate musically elite subjects from those subjects less musically elite. The experiment was undertaken in an effort to evaluate an objective measure which Iowa high school teachers

¹Raleigh M. Drake, Manual for the Drake Musical Aptitude Tests (Chicago: Science Research Associates, 1957), p. 16.

could use to assist them in the identification of musically talented students. The investigation was designed to employ a criterion for validity which would reflect not only teacher estimates of a subject's musical characteristics but also the student's interests and specific participation in music activities.

II. MUSICAL APTITUDE TESTS: CRITERIA OF VALIDITY

Various criteria have been used to determine the validity of existing aptitude measures. With regard to the validity of the Measures of Musical Talent, Seashore has stated that

... they have been validated for what they purport to measure. This is an internal validation in terms of success in the isolation of the factor measured and the degree of control of all other factors in the measurement.²

Seashore went on to state that the tests of the battery

... should not be validated in terms of their showing on an omnibus theory or blanket rating against all musical behavior, including such diverse and largely unrelated situations as composition, directing, voice, piano, violin, saxophone, theory, administration, or drums; because there are hundreds of other factors which help to determine job analysis in each of such fields.³

²Carl E. Seashore, "Two Types of Attitude Toward the Evaluation of Musical Talent," Music Educators Journal, XXIV (December, 1937), p. 25.

³Ibid.

The unwillingness of many investigators to accept Seashore's philosophy of test validation is shown by the number of independent studies which have been conducted upon the Seashore battery by Mursell,⁴ Taylor,⁵ Drake,⁶ and others. Most of these investigations used success in applied music or grades in various music courses as criteria. Mursell, a leading exponent of the omnibus theory, contended that there was but one method of validating a test of musical aptitude.

... and that is to ascertain whether persons rating high or low or medium on these tests also rate high and low and medium in what one may call "musical behavior," i.e., sight singing, playing the piano, getting through courses in theory and applied music, and the like.⁷

Validity coefficients determined by investigators

⁴James L. Mursell, "Measuring Musical Ability and Achievement: A Study of the Correlation of the Seashore Test Scores and other Variables," Journal of Educational Research, XXVI (1932), pp. 116-126.

⁵Elizabeth M. Taylor, "A Study of the Prognosis of Musical Talent," Journal of Experimental Education, X (1941), pp. 1-28.

⁶Raleigh M. Drake, "The Validity and Reliability of Tests of Musical Talent," Journal of Applied Psychology, XVII (1933), pp. 447-58.

⁷James L. Mursell, "What About Music Tests," Music Educators Journal, XXIV (1937), p. 16.

of the Seashore test who used the criteria described by Mursell were usually low.⁸ These validity coefficients could in part be attributed to the tests' low reliability coefficients. Seashore accounted for this poor reliability by the limitations imposed upon the testing device:

... It must be remembered that the phonograph records are a makeshift for the purpose of securing a dragnet group test of an unselected population in a limited period of time and without training for observation.⁹

Schoen reported the development of a test battery in 1923.¹⁰ He investigated its validity by using a criterion of teacher estimates. A student's ability in tone production, tone inflection, phrasing, and talent as a whole were rated. The results of Schoen's investigation were presented in the form of talent charts. He employed only ten subjects for the validation of his test. It is the writer's opinion that such limited investigation is inadequate.

The Kwalwasser-Dykema music tests were published in 1930.¹¹ No reliability or validity data is reported in the

⁸Robert W. Lundin, An Objective Psychology of Music (New York: Ronald Press Company, 1953), p. 208.

⁹Seashore, op. cit., p. 26.

¹⁰Max Schoen, "The Validity of Tests of Musical Talent," Journal of Comparative Psychology, III (1923), pp. 101-121.

¹¹Jacob Kwalwasser and Peter Dykema, Kwalwasser-Dykema Music Tests (New York: Carl Fischer, Inc., 1930).

test manual. However, as a result of an independent investigation of the validity of the Kwalwasser-Dykema battery designed to use a criterion of teacher estimates, Farnsworth has concluded that the tests are "... too unreliable for purposes of clinical diagnosis."¹²

The Tilson-Gretsch Test for Musical Aptitude was published in 1939. Tilson investigated the validity of the tests by comparing the results with two separate criteria; the grades received by the subjects in music classes in elementary and secondary schools and their scores on the Seashore Measures of Musical Talent.¹³ Tilson reported that the results indicated that the Tilson-Gretsch battery was slightly superior to the Seashore tests for the fifth through the eighth grades. However, Farnsworth pointed out that reliability coefficients for college students and adults were considerably lower than the reliabilities in parallel Seashore tests.¹⁴ On the basis of his investigation, Farnsworth concluded that the Tilson-Gretsch battery

¹²Paul R. Farnsworth, "An Historical, Critical, and Experimental Study of the Seashore-Kwalwasser Test Battery," Genetic Psychology Monograph, IX (1931), p. 384.

¹³L. M. Tilson, "A Study of the Prognostic Value of the Tilson-Gretsch Test for Musical Aptitude," Teachers College Journal, XII, No. 5 (1941), p. 110.

¹⁴Paul R. Farnsworth, "Data on the Tilson-Gretsch Test for Musical Aptitude," Journal of Musicology, IV, No. 2 (1945), p. 101.