A STUDY OF THE VALIDITY OF THE IOWA TESTS OF MUSIC LITERACY

by

James L. Mohatt

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Music in the Graduate College of The University of Iowa

January, 1971

Thesis Chairman: Professor Edwin Gordon
Co-Chairman: Associate Professor Darrell L. Sabers

Music education has reached a period of maturity and development in which educators are seeking the establishment of common objectives in music instruction, and also effective and efficient means of evaluating students' progress toward achieving accepted goals. Gordon, after identifying music literacy as the main objective of early instruction in music, has constructed a multi-level battery of music achievement tests, the Iowa Tests of Music Literacy (ITML). Each of the six levels of ITML contains six subtests: Tonal Concepts, Aural Perception (T-1); Tonal Concepts, Reading Recognition (T-2); Tonal Concepts, Notational Understanding (T-3); Rhythmic Concepts, Aural Perception (R-1); Rhythmic Concepts, Reading Recognition (R-2); and Rhythmic Concepts, Notational Understanding (R-3).

The purpose of this study was to investigate the content and criterion-related validity of the ITML battery. Specifically, the first problem was to examine the content validity and experimentally determine whether the six different test levels actually comprise items which become sequentially more difficult. And the second problem was to investigate the criterion-related validity by comparing students' subtest and composite scores with various measures of their demonstrated musical behaviors.

First, all six levels of ITML were administered to an entire eighth grade of one public school. Means,

standard deviations, reliabilities, score distributions, and intercorrelations within and between levels were examined for purposes of investigating content validity of ITML.

Second, those students who completed all levels of ITML were asked to demonstrate specific musical understandings and behaviors which were objectively evaluated. ITML scores were then correlated with these validity criteria in order to investigate the criterion-related validity of the tests.

ITML reliabilities were found to be generally in the .60's and .70's for subtests, in the .70's and .80's for total tests, and in the .80's and .90's for composite tests for all levels. The means for three of the six subtests (T-2, R-1, and R-2) do generally and sequentially decrease as the levels increase numerically; but the means of three of the subtests (T-1, T-3, and R-3) do not do so. The score frequency distributions confirm the findings reported for the means.

Intercorrelations within levels indicate that the subtests are generally interrelated to a rather low degree, from about .20 to about .40. Intercorrelations between levels indicate that the subtests are related from level to level to a comparatively high degree, from about .30 to about .50. And the composite intercorrelations indicate that all levels correlate with each other at about .70.

The investigation of the criterion-related validity

of ITML indicated that subtests correlate with corresponding validity criteria to a rather high degree, from about .30 to about .60. And composite ITML scores correlate to a very high degree with the composite validity criteria, from about .60 to about .70.

On the basis of the results, the content validity of ITML can be only partially objectively substantiated. Means, standard deviations, reliabilities, and intercorrelations among the tests within and between levels indicate that the battery is appropriate for use with public school students. But only three of the six subtests could be shown to become systematically more difficult.

The investigation of the criterion-related validity indicated that ITML does have a very high degree of criterion-related validity and that the tests can serve as a valuable tool for efficiently and objectively measuring and evaluating students' progress in music literacy skills.

Abstract approved: _____, Thesis supervisor

_____, title and department

_____, date

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

CHAPTER I

PURPOSE OF THE STUDY

Introduction

"The primary function of any evaluation procedure is to determine to what extent students have achieved the objectives of instruction."[1] Measurement and evaluation of achievement in elementary and secondary music education are necessary parts of the learning process, not only as a means of determining rate and degree of progress, but also as a means of identifying the objectives of instruction. That which teachers emphasize in the evaluation of pupils, especially in objective evaluation through the use of achievement tests, defines to students what teachers consider important. Lehman states that achievement tests in music have not been generally accepted by teachers because there is no agreement on specific outcomes that should be expected as a result of instruction in music.[2]

Because elementary and secondary music educators

---

1. Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education, second edition, (New York: John Wiley and Sons, 1961). p. 29.

2. Paul R. Lehman, Tests and Measurements in Music, (Englewood Cliffs, New Jersey: Prentice-Hall, 1968). p. 57.

attempt, in their teaching efforts, to include types of learning which are difficult to measure, unanimity in identifying and determining priorities of objectives for music instruction has, in the past, been more difficult than in most other subject-matter areas of the school curriculum. Appreciation for and receptive attitude toward a variety of styles of music from all periods of history are commonly considered to be desirable outcomes of music instruction. But evidence of improvement or growth in appreciation and attitude has been difficult to measure. Also, the degree of relationship between music instruction and any indicated improvement or growth in appreciation and attitude, although probably present to some extent, does not easily lend itself to objective measurement.

The non-academic or extracurricular status of some music instruction and the concessions of time, effort, and importance afforded some prestigious performance groups in schools probably have played a part in retarding the development of uniform minimum standards of musical understanding and achievement. Recently, governmental and other agencies have been exerting increased pressure for effective evaluation of all educational programs. As state and federal aid to schools increases it is likely that more and better evaluation will help to determine the effectiveness of the tax dollars spent. Those school programs which cannot

produce measurable results may suffer reduced support.

According to Reimer:

> The ability of research to give precise knowledge about quantifiable behaviors is just as great for us as it is in any other field. Since many aspects of the teaching and learning of music are inherently objective, we are obliged, as are all teachers and scholars, to seek verifiable truth in such matters wherein such truth exists.[3]

And Van Bodegraven says that:

> . . . . there is a growing conviction that the program of music education has matured to a point where there should be a core of specific and measurable outcomes of experiences in all phases of music and that these outcomes should be common for all students . . . .[4]

It is apparent that there is a need for identifying and defining desired outcomes of music instruction. And there is also a need for objectively measuring and evaluating achievement of the desired outcomes. Gordon has identified and defined music literacy as one of the main objectives of the early years of music instruction, and he has also described it as achievement which can be objectively measured and evaluated. He has indicated the

---

3. Bennett Reimer, "Effects of Music Education," Perspectives in Music Education, Source Book III. Washington, D.C.: Music Educators National Conference, 1966. p. 473.

4. Paul Van Bodegraven, "Music Education in Transition," Perspectives in Music Education, Source Book III. Washington, D.C.: Music Educators National Conference, 1966. p. 30.

manner and the order in which children acquire basic musical concepts and skills, and also the means by which teachers are able to induce, first, the understandings necessary for music literacy readiness, and then music literacy itself. The following excerpt contains his interpretation of music literacy.

> . . . . students become truly musically literate because they first develop tonal and rhythmic aural perception. These musical reactions are then simply associated with corresponding notational symbols for the purpose of music literacy comprehension. Student achievement in these abilities - aural and kinesthetic perception plus music literacy - which together constitute music listening readiness, can be objectively evaluated to, among other important purposes, guide teachers in the instruction of more complex aspects of music appreciation.[5]

The Iowa Tests of Music Literacy[6] (ITML), developed by Gordon, is a battery of six levels of tape-recorded music achievement tests which measure aural perception, reading recognition, and notational understanding of tonal concepts and rhythmic concepts as interpreted by him. The six levels of ITML answer sheet booklets may

---

5. Edwin Gordon, Psychology of Music Teaching, (Englewood Cliffs, New Jersey: to be published by Prentice-Hall in 1971).

6. Edwin Gordon, Iowa Tests of Music Literacy, (Iowa City: Bureau of Educational Research and Service, Division of Extension and University Services, The University of Iowa, 1970).

be found in Appendix A.

The subtests of all levels of ITML, although designed to measure parallel skills, progress in difficulty from Level 1 through Level 6. The organization of the six levels is similar but the content is different.

Purpose of the Study

ITML has been designed to determine the extent to which students develop aural perception and music literacy, the main objectives of early music instruction. In order for ITML to be useful and practical in schools, the battery must be proven to function as a valid instrument in measuring that which it is intended to measure. The purpose of this study was to investigate the content and criterion-related validity of the ITML battery.

Problems of the Study

Specifically, the first problem of the study was to examine the content validity of ITML and experimentally determine whether the six different test levels actually comprise items which become sequentially more difficult.

The second problem of the study was to investigate the criterion-related validity of the battery of tests by comparing students' subtest and composite scores with various measures of their demonstrated musical behaviors.

Description of the Iowa Tests of Music Literacy

The six levels of ITML are divided into two aspects of fundamental musical skills, Tonal Concepts and Rhythmic Concepts. Each of these tests includes three subtests: Aural Perception, Reading Recognition, and Notational Understanding. All of the test items were tape-recorded on the Moog Synthesizer by Professor Peter Lewis, a member of the music faculty of The University of Iowa.

The first Tonal Concepts subtest, Aural Perception (T-1), requires the student to identify, by filling ovals, the recorded music phrases as being, depending on the specific level of the test, either in major or minor, or in a usual or an unusual mode. There are twenty-two items in this subtest at each level.

The second Tonal Concepts subtest, Reading Recognition (T-2), also consists of twenty-two items for each level. The student is required, on all levels, to indicate by filling ovals whether the phrases he hears on the tape recording are the same as those notated on the answer sheet.

The third Tonal Concepts subtest, Notational Understanding (T-3), contains, depending on the level, six, seven, or eight items. The student is required to fill one of two blank note heads to indicate the correct pitch for five of the nine notes heard in each phrase.

The first Rhythmic Concepts subtest, Aural Perception

(R-1), requires the student to identify, by filling ovals, the recorded rhythmic patterns as being, depending on the specific level of the test, in either duple or triple, usual or mixed, usual or unusual, mixed or unusual meter. There are twenty-two items in this subtest at each level.

The second Rhythmic Concepts subtest, Reading Recognition (R-2), also consists of twenty-two items for each level. The student is required, on all levels, to indicate by filling ovals whether the rhythm patterns he hears on the tape recording are the same as those notated on the answer sheet.

The third Rhythmic Concepts subtest, Notational Understanding (R-3), contains, depending on the level, nine, ten, eleven, or twelve items, each of which is a rhythmic phrase. The student is required to fill blank note heads, beams, and rests in order to indicate, in correct rhythmic notation, the phrase heard on the recording.

Each subtest is prefaced by recorded directions and practice exercises. All of the subtests except those of Notational Understanding have question mark ovals so that students need not guess answers. The three Tonal Concepts tests of each level require approximately forty-five minutes of administrative time, as do the three Rhythmic Concepts tests.

CHAPTER II

RELATED LITERATURE

Introduction

Literature related to the development, construction, and functions of multi-level music literacy achievement tests is almost nonexistent. Only one experimental study, "An Investigation of the Adequacy of the Content and Difficulty Levels of the Iowa Tests of Music Literacy,"[7] conducted by Warren Swindell, is closely related to the present study.

The Swindell Study

The purpose of Swindell's investigation was to determine, first, whether the content of ITML is appropriate for public school students in Grades 4 through 9 and, second, whether the original test level classifications are in the proper difficulty hierarchy. He administered ITML to a homogeneous socio-economic population of students in Grades 4 through 9. All students in each classroom who participated in the study were administered

---

7. Warren Swindell, "An Investigation of the Adequacy of the Content and Difficulty Levels of the Iowa Tests of Music Literacy," (unpublished Ph.D. thesis, The University of Iowa, 1970).

the tests, but each student was given only one test level. Random order of test administration was followed in a manner which provided for each level of the tests to be given in every grade.

The means, standard deviations, and reliabilities for each of the six levels of ITML for Grades 4 through 9 may be found in Appendix B, Tables 47-52. However, it should be noted that the Notational Understanding subtests in both the Tonal Concepts and Rhythmic Concepts divisions differ in total possible scores from level to level. As a result, a comparison of means as percentages of possible total scores of subtests within each level and from level to level is more useful for the purposes of this study than is a comparison of means as raw scores. Therefore, the investigator has converted all subtest and composite mean scores reported by Swindell to percentages of possible total scores. These converted means are reported in Tables 1-7.

The subtest means analyzed as a percentage of the total possible scores indicate that, for Grades 7-9, Tonal Concepts, Aural Perception, Level 4; Tonal Concepts, Reading Recognition, Level 2; and Rhythmic Concepts, Aural Perception, Level 3, yield higher scores than do the corresponding subtests of lower levels in the battery. And for Grades 4-9, Tonal Concepts, Notational

TABLE 1

(Swindell Study)

ITML SUBTEST MEANS CONVERTED TO A PERCENTAGE OF TOTAL POSSIBLE - GRADE 4

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Subtest T-1	60	51	46	53	48	50
Subtest T-2	70	64	61	53	48	48
Subtest T-3	65	57	66	56	43	53
Subtest R-1	59	53	54	43	46	43
Subtest R-2	54	56	54	47	46	54
Subtest R-3	60	54	58	57	61	56