

A COMPARATIVE STUDY OF IOWA TESTS OF MUSIC LITERACY  
NORMS FOR UNITED STATES DEPENDENT STUDENTS  
ATTENDING SCHOOL IN FRANKFURT, GERMANY,  
WITH NORMS DERIVED FOR UNITED STATES  
STUDENTS WHO PARTICIPATED IN THE  
STANDARDIZATION OF THE TEST  
BATTERY

by

Robert Lee Romine

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

July, 1973

Thesis supervisor: Professor Edwin Gordon

## ABSTRACT

The specific problem of this study was to compare Iowa Tests of Music Literacy norms for students in United States Dependents Schools, European Area, in Germany with those developed for students who participated in the ITML national standardization program. Ancillary problems of the study were to: 1) investigate the test-retest reliability of each ITML subtest, the Total Tonal Concepts and Rhythmic Concepts tests, and the Composite test score, 2) investigate the optimum length of each ITML Level One subtest, 3) investigate the appropriateness of the Spearman-Brown Prophecy Formula for use with ITML Level One, 4) investigate the practice effects of taking ITML a second time, and 5) compare ITML norms for native German students who understand the English language with those developed for students who participated in the ITML national standardization program.

During the 1971-72 school year 843 American students enrolled in two USDESEA schools in Frankfurt, Germany, were tested and then retested on ITML. This total included 292 elementary students and 551 students enrolled in junior high school. An additional 452 native German students were given a single administration of the test battery for an ancillary problem. Of this total, 157 students were enrolled in grades four through six, 139 students in grades seven through nine, and 127 students in grades ten through twelve.

In order to compare scores of the first administration of ITML of USDESEA students and the single administration of German students with the scores of the American students, the nine standard score distributions were plotted and smoothed graphically to determine percentile ranks for each group of students. The percentile ranks chosen for comparison were: 99, 95, 90, 75, 50, 25, 10, 5, and 1. For purposes of this study, consistent differences (within a distribution) of five standard score points

or more at the selected percentile ranks were considered to represent consequential normative discrepancies.

Of the eighteen normative comparisons at selected standard score-percentile rank equivalents for the USDESEA and American elementary students, only two subtests yielded consequential discrepancies. The Rhythmic Notational Understanding subtest for Level Two favored the USDESEA students, whereas the Tonal Aural Perception subtest for Level Three favored the American students.

Of the 36 comparisons for junior high school students, ten subtests were found to have consequential normative discrepancies. The discrepancies on Rhythmic Notational Understanding subtests for Levels Two and Five, and the Tonal Notational Understanding subtests for Level Four and Level Five favored the USDESEA students. On the Rhythmic Reading Recognition subtest for Level Four, both USDESEA and American students were favored at selected ranges. The five subtests which favored the American students were the Rhythmic Aural Perception for Level Two, the Tonal

Reading Recognition for Level Four, the Rhythmic Notational Understanding for Level One and Level Six, and the Rhythmic Notational Understanding subtest for Level Four.

For the native German students, consequential normative discrepancies were found for 16 of the 36 subtests. Of this total, five were for the elementary students, five were for the junior high school students, and six were for the high school students.

As expected, the test-retest reliabilities were lower than the published split-halves reliabilities derived from the ITML standardization program. Although subtests for both USDESEA groups had lower reliabilities, the reliabilities for the USDESEA junior high students were closer to the established norms than were those for the elementary students.

Concerning the optimum length of the ITML subtests, it was found with one exception, that reliability is highest for the total length of each subtest.

Through the use of the Spearman-Brown Prophecy Formula, the predicted reliability coefficients were generally more accurate as the subtest sections increased in length.

The second administration observed means were generally higher than those derived from the first administration for all levels for both USDESEA groups. The standard score means were significantly higher for the second administration for six subtests in the elementary school grade range and for eight subtests in the junior high school grade range.

In general, the norms provided in the ITML Manual for all six ITML levels can be used with USDESEA students. However, results on the following subtests should be interpreted with caution:

Elementary Students

- 1) Level Two - Rhythmic Notational Understanding
- 2) Level Three - Tonal Aural Perception

Junior High Students

- 1) Level One - Rhythmic Reading Recognition
- 2) Level Two - Rhythmic Aural Perception and Rhythmic Notational Understanding
- 3) Level Four - Tonal Reading Recognition, Tonal Notational Understanding, Rhythmic Reading Recognition, and Rhythmic Notational Understanding

- 4) Level Five - Tonal Notational Understanding  
and Rhythmic Notational Under-  
standing
- 5) Level Six - Rhythmic Reading Recognition

Regarding the optimum length of the ITML subtests, it is possible that they should be lengthened to ensure greater test stability.

Abstract approved: Edu. Lohr, Thesis supervisor  
Prof. Music title and department  
July 10, 1973 date

## TABLE OF CONTENTS

CHAPTER	Page
	v
	1
I	1
Introduction.....	3
Purpose of the Study.....	4
Problems of the Study.....	5
Description of the <u>Iowa Tests of Music Literacy</u> .....	9
Footnotes for Chapter I.....	10
II	10
Introduction.....	10
The Brown Study.....	10
The Gordon Study.....	12
Comparison of Related Studies with Present Study.....	19
Footnotes for Chapter II.....	20
III	21
Subjects.....	21
Testing Procedures.....	22
Data Analysis.....	25
Footnotes for Chapter III.....	33
IV	34
Standard Score Means, Standard Deviations, and Standard Errors of Measurement.....	34
Test-Retest Reliability.....	46
Optimum Length of Subtests.....	50



TABLE OF CONTENTS (Cont.)

CHAPTER	Page
Appropriateness of the Spearman-Brown Prophecy Formula.....	53
Practice Effects of Taking <u>ITML</u> .....	56
Intercorrelations Among the Subtests....	67
Norms for USDESEA Students.....	67
Norms for German Students.....	89
V SUMMARY AND CONCLUSIONS.....	97
Problems of the Study.....	97
Design.....	98
Results.....	100
Conclusions.....	103
BIBLIOGRAPHY.....	104

## CHAPTER I

## PURPOSE OF THE STUDY

Introduction

During the last decade the American society has placed increased emphasis on the rights and needs of the individual with respect to educational opportunities. For purposes of evaluating learning that has taken place through courses of study designed to meet students' individual needs, educators have come to rely upon standardized achievement tests. However, in comparison to other academic areas of achievement, there are only a few standardized music tests. Colwell, in his description of published music tests, states that music "achievement tests are almost nonexistent."<sup>1</sup>

A relatively new music achievement battery is the Iowa Tests of Music Literacy,<sup>2</sup> which was standardized during the 1970-71 school year.

ITML,\* a six-level battery of music achievement

---

\*Hereafter the Iowa Tests of Music Literacy will be titled in abbreviated form, ITML.

tests, is designed to measure students' progress in tonal and rhythmic aural perception and music literacy achievement. The results of the ITML subtests may be used to: 1) diagnose a student's individual musical strengths and weaknesses, 2) compare a student's musical achievement to his musical potential as indicated by Musical Aptitude Profile results, 3) evaluate the extent of a student's continuous musical development, and 4) compare a student's relative standing in musical achievement among other students.<sup>4</sup>

Mohatt, in his investigation of the validity of ITML, found the battery to "have a very high degree of criterion-related validity."<sup>5</sup> Swindell concluded that the content of the battery was appropriate for public school students in grades 4 through 9.<sup>6</sup> And, individual researchers, such as Schleuter<sup>7</sup> and Thayer,<sup>8</sup> found that although the relationship between MAP\* and ITML is substantial, the batteries

---

\*MAP is an abbreviation for the Musical Aptitude Profile.

do serve the specific purposes for which each was designed.

#### Purpose of the Study

Users of standardized tests assume that the established norms derived from a standardization program can be used with a high degree of confidence because: 1) the tests are usually constructed by experts in the field, 2) through revisions, the tests are strengthened by the deletion of weak and non-discriminating test items, and 3) the tests are expected to be more reliable and valid than teacher-made tests.<sup>9</sup>

Test administrators teaching in United States Dependents Schools, however, may question the appropriateness of published norms for evaluating their students because of differences in cultural background and environment. The students attending these schools, as a result of their fathers' occupation, will probably have attended a minimum of three different schools in at least two foreign countries before graduating from high school. Also, most of these students will not have had the

opportunity to study instrumental music until entering junior high school due to the limited curriculum offered in most elementary Dependents Schools. Therefore, the main purpose of this study was to investigate the appropriateness of the norms established in the national standardization of ITML for use with American dependent students who attend the USDESEA\* schools in Germany.

#### Problems of the Study

The specific problem of this study was to compare ITML norms for students in USDESEA schools in Germany with those developed for students who participated in the ITML national standardization program.

Ancillary problems which related to the study were to:

- 1) Investigate the test-retest reliability of each ITML subtest, the total Tonal Concepts and Rhythmic Concepts tests, and the Composite test score.
- 2) Investigate the optimum length of each ITML Level One subtest.
- 3) Investigate the appropriateness of the Spearman-

---

\* USDESEA is an abbreviation for United States Dependents Schools, European Area.

Brown Prophecy Formula for use with ITML Level One.

4) Investigate the practice effects of taking ITML a second time.

5) Compare ITML norms for native German students who understand the English language with those developed for students who participated in the ITML national standardization program.

#### Description of the Iowa Tests of Music Literacy

The Iowa Tests of Music Literacy battery is comprised of six levels that become sequentially more complex from Level One through Level Six. Each of the six levels of ITML is divided into two sections: Tonal Concepts and Rhythmic Concepts. Each of these sections includes three subtests: Aural Perception, Reading Recognition, and Notational Understanding. All test items are tape recorded and performed on the Moog Synthesizer. Recorded directions and test examples precede each subtest.

There are 22 test items on each level of the Tonal and Rhythmic Aural Perception and Reading Recognition subtests. For each of these subtests the student is required to fill one of three ovals

indicating which answer he believes to be correct; the third oval is an "in doubt" response. Depending on the level of the test, the Tonal Notational Understanding subtest has from six to eight test items and the Rhythmic Notational Understanding subtest has from nine to fourteen test items. For each of these subtests the student is asked to complete either the tonal or rhythmic notation on the answer sheet-booklet.

The items in the Tonal Aural Perception subtests (T-1) are answered by filling the oval which indicates whether the seven tones heard, depending on the level, are in major or minor tonality, or are "usual" or "unusual."\* For the Tonal Reading Recognition subtests (T-2) the student is required to indicate, by filling an oval marked yes (Y) or

---

\*"The terms 'major' and 'minor', like those of 'modal' and 'unusual', are referred to as type of tonality. However, in this connection, although major and minor correspond to the Ionian and Aeolian modes, respectively, from practical usage, they are referred to as usual modes. Dorian, Phrygian, Lydian, Mixolydian, and Locrian, because of their less frequent use, are referred to as unusual modes." Gordon, ITML Manual, p. 9.

no (N), if the seven-tone phrases he hears on the tape recording are the same as those notated on the answer sheet-booklet. The third Tonal Concepts subtest, Notational Understanding (T-3), consists of a recorded nine-tone pattern for each test item. Notes are given on the answer sheet-booklet for four of the nine tones heard on the recording. The student is required to fill blank note heads for each of the five remaining tones he hears.

For the first Rhythmic Concepts subtest Aural Perception (R-1), the student indicates if the recorded test item is in either duple or triple, usual or mixed, usual or unusual, mixed or unusual meter, depending on the level of the test.\* For the Rhythmic Reading Recognition subtest (R-2),

---

\*"The terms 'duple' and 'triple', like those of 'mixed' and 'unusual', are referred to as types of meter. Patterns written in duple meter are those which generally have a 2 or 4 as the upper numerals in the meter signature. Patterns written in triple meter are those which generally have a 3, 6, or 12 as the upper numeral of the meter signature. Mixed meter patterns include triplets or duplets, but have upper numerals in the meter signature which designate duple meter or triple meter, respectively. And, patterns written in unusual meter are generally those which have a 5, 7, 9, or 11, and sometimes a 3 as the upper numerals of the meter signature. Both duple and triple are considered usual meter as compared to mixed and unusual meter." Gordon, ITML Manual, p. 10.



the student is required to fill the ovals marked yes (Y) or no (N), to indicate if what he hears on the tape recording is the same as the rhythm notation on the answer sheet-booklet. For the Rhythmic Notational Understanding subtest (R-3), the student is required to fill blank note heads, flags, beams, ties, and/or rests to complete the partially written rhythm patterns on the answer sheet-booklet.

The Iowa Tests of Music Literacy were standardized during the 1970-71 school year. A total of 18,680 students, enrolled in 27 school systems in 13 states, grades four through twelve, were included in the norms sample. Percentile norms were established for each ITML level for three grade ranges: 1) elementary, grades four through six, 2) junior high school, grades seven through nine; and 3) high school, grades ten through twelve.<sup>10</sup>

Footnotes for Chapter I

<sup>1</sup>Richard Colwell, The Evaluation of Music Teaching and Learning (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970), p. 144.

<sup>2</sup>Edwin Gordon, Iowa Tests of Music Literacy (Iowa City, Iowa: Bureau of Educational Research and Service, The University of Iowa, 1970).

<sup>3</sup>Edwin Gordon, Musical Aptitude Profile (Boston: Houghton Mifflin Company, 1965).

<sup>4</sup>Edwin Gordon, Iowa Tests of Music Literacy Manual (Iowa City, Iowa: Bureau of Educational Research and Service, The University of Iowa, 1970), pp. 1-2.

<sup>5</sup>James L. Mohatt, "A Study of the Validity of the Iowa Tests of Music Literacy," (Ph.D. dissertation, The University of Iowa, 1971), p. 84.

<sup>6</sup>Warren C. Swindell, "An Investigation of the Adequacy of the Content and Difficulty Levels of the Iowa Tests of Music Literacy," (Ph.D. dissertation, The University of Iowa, 1970), p. 137.

<sup>7</sup>Stanley Schleuter, "A Study of the Inter-relationship of Personality Traits, Musical Aptitude, and Musical Achievement," (Ph.D. dissertation, The University of Iowa, 1971).

<sup>8</sup>Robert Thayer, "An Investigation of the Interrelation of Personality Traits, Musical Achievement, and Different Measures of Musical Aptitude," (Ph.D. dissertation, The University of Iowa, 1971).

<sup>9</sup>Paul Lehman, Tests and Measurements in Music (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968), p. 29.

<sup>10</sup>Gordon, ITML Manual, pp. 95-96.

## CHAPTER II

### REVIEW OF RELATED STUDIES

#### Introduction

Previous investigations of the musical aptitude test battery, MAP, by Brown<sup>1</sup> and Gordon,<sup>2</sup> were similar to the present investigation. In both studies, norms which were derived for students in unique school systems were compared with those derived from the national standardization program.

#### The Brown Study

Although the major problem of Brown's study was to investigate the optimum length of each MAP subtest, an ancillary problem was to compare score distributions for parochial school students with the established national standardization data. Brown tested a total of 896 students in Catholic schools in Kenosha, Wisconsin. Of this total, 331 were fifth grade students, 307 were seventh grade students, and 258 were ninth grade students.