AN INVESTIGATION OF THE COMPARABILITY OF AMERICAN AND GERMAN NORMS FOR THE MUSICAL APTITUDE PROFILE

by

Arthur William Schoenoff

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

December, 1972

Thesis Supervisor: Professor Edwin Gordon

TABLE OF CONTENTS

CHAPTER I

PURPOSE OF THE STUDY

Introduction

Experimental investigations in which skills of students from different nationalities were compared can be traced back almost a century. Myers reported that in 1879 Herzen investigated the physical reaction-time of Italians and Germans. "His unpublished experiments led him [Herzen] to believe that Italians react more slowly than Germans."[1] In regard to musical ability, Myers conducted experiments in 1888-1889 with natives of Murray Island, which is located in the Torres Straits between Australia and Papua. In an investigation to determine the highest pitch perceptible by Murray Island natives, as compared to subjects in Aberdeenshire, Scotland, Myers blew a Galton-whistle which produces sounds of varying pitches by changing the length of its air column.[2] Myers

---

1. Charles S. Myers, "Reaction-times," Reports of the Cambridge Anthropological Expedition to the Torres Straits, 6 vols., ed. A. C. Haddon (Cambridge, England: University Press, 1903), II, 221.

2. Myers, "Upper limit of hearing," op. cit., 149-154.

concluded that (1) Murray Island subjects and Aberdeenshire subjects were nearly identical in the upper limit of their hearing, and (2) the children of both groups could hear higher pitches than the adults. Also during his expedition to Murray Island, Myers investigated the ability of native subjects to perceive small pitch differences sounded by two 256 Hz[3] tuning forks, one of which had attached to it a sliding metal bar which caused the sounding-pitch of the fork to be altered.[4] Holding each fork about 15 cm from the ear of a subject, Myers struck each of the two forks in close sequence; each fork was permitted to sound for about two seconds. Again comparing Murray Island subjects to Aberdeenshire subjects, Myers concluded that the Aberdeenshire subjects were capable of perceiving smaller pitch differences than were the Murray Island subjects. However, Myers observed that there was less difference in this ability between the children of the two groups of subjects.

Subsequent to these early studies, more than fifty investigations were designed to compare musical ability among sample populations which were dissimilar in some

---

3. Hz is the symbol for "hertz," a unit of frequency equal to one cycle per second.

4. Myers, "The smallest perceptible tone-difference," op. cit., 155-168.

social or economic aspect other than age or sex. These investigations were given impetus by the work of Carl Seashore. Early in the present century at the University of Iowa, Seashore began developing standardized tests which were designed to measure musical aptitude. This work culminated in 1919 with the first version of the Seashore Measures of Musical Talent, published and recorded on phonograph discs.[5] With this battery, the music psychologist had available for the first time a standardized test with which to conduct investigations of musical aptitude.

Murdock was the first person to report the use of the 1919 Seashore battery to investigate differences in musical aptitude between groups of subjects who had different national ancestries. In 1923 Murdock administered part of the Seashore battery to Chinese, Japanese, Korean, Hawaiian, Chinese-Hawaiian, Portuguese, and Anglo Saxon-Hawaiian subjects living in the Hawaiian Islands.[6] Only data from the pitch test were analyzed. The test results were reported in terms of the percentage of subjects who made higher scores than the score given for the median on

---

5. Carl E. Seashore, Seashore Measures of Musical Talent (Chicago: C. H. Stoelting Company, 1919).

6. Katherine Murdock, "A study of differences found between races in intellect and morality, II," School and Society, XXII/569 (November 21, 1925), 659-664.

the published norms. Only the Chinese-Hawaiians exceeded fifty percent in this statistic.

There have been ten studies reported which were based on the administration of a major portion of a standardized musical aptitude test to an atypical sample of at least five hundred subjects. These studies were reported by Ausherman,[7] Brown,[8] Candor,[9] Dykema,[10] Garth and Isbell,[11] Gordon,[12] Johnson,[13] Sanderson,[14] Wilson,[15] and

---

7. Paul F. Ausherman, "The establishing of Kwalwasser-Dykema test norms for the Negro race" (unpublished Master's thesis, Indiana State University, Terre Haute, 1936).

8. Merrill E. Brown, "An investigation of the optimum length of Musical Aptitude Profile subtests" (unpublished doctoral dissertation, University of Iowa, Iowa City, 1967).

9. Ethel Candor, "The musical talent of Mexican school children" (unpublished Master's thesis, University of Denver, Colorado, 1933).

10. Peter W. Dykema, "An international study of music talent," Music Educators National Conference; Thirtieth Yearbook (1937), 94-96.

11. Thomas R. Garth and Sarah R. Isbell, "The musical talent of Indians," Music Supervisors Journal, XV/3 (February, 1929), 83-87.

12. Edwin Gordon, "A comparison of the performance of culturally disadvantaged students with that of culturally heterogeneous students on the Musical Aptitude Profile," Psychology in the Schools, IV/3 (July, 1967), 260-262.

13. Guy B. Johnson, "A summary of Negro scores on the Seashore music talent tests," The Journal of Comparative Psychology, XI/4 (April, 1931), 383-393.

Witherson.[16] However, in a majority of these studies, only the mean or median test scores were used to compare sample group norms to published norms. In all, only five studies were found which were designed to compare overall sample norms to published norms. These five studies are described in detail in Chapter II.

Purpose of the Study

There is a growing desire in the Federal Republic of Germany for sound psychological testing in the schools. Concurrent with this desire is an interest in tests developed in America.[17] Regarding musical aptitude tests, none

---

14. Helen E. Sanderson, "Differences in musical ability in children of different national and racial origin," Journal of Genetic Psychology, XLII (1933), 100-119.

15. Margaret W. Wilson, "A comparison of scores earned by one thousand Negro and white children on the Kwalwasser-Dykema Music Tests (unpublished Master's thesis, Syracuse University, New York, 1935).

16. Maude I. Witherson, "The effects of race and nationality upon music measurement" (unpublished Master's thesis, Syracuse University, New York, 1935).

17. These opinions are based on correspondence between an official for a German publisher (Westermann-Verlag), a German "music scientist and sound engineer" (Dr. Kleinen), and the author of the Musical Aptitude Profile, plus personal conversations between the writer and German educators. A similar view regarding testing was reported in: "Education and science," Kulturbrief, ed. Hildegard Brodda, XI/5 (May, 1972), 19.

of the more popular currently published batteries have been reported to be standardized for use outside the country of their origin. However, two batteries, those of Seashore[18] and Bentley,[19] have been translated for use with German-speaking subjects, although neither battery has been standardized on German students.

Because of its thorough preparation and demonstrated high experimental validity, specific interest in the Musical Aptitude Profile[20] has been expressed by German educators to both the test author and the writer. However, if a psychological test is to be administered for evaluative purposes to a particular population, the adequacy of the test for use with that population should be investigated. Therefore it is the purpose of this study to investigate the appropriateness of MAP for a representative sample of native German students who are administered the test battery with tape recorded translated directions.

---

18. Carl E. Seashore, Don Lewis, and Joseph G. Saetveit, Seashore Measures of Musical Talents (New York: The Psychological Corporation, 1960).

19. Arnold Bentley, Measures of Musical Ability (New York: October House, Inc., 1966).

20. Edwin Gordon, Musical Aptitude Profile (Boston: Houghton Mifflin Company, 1965). Subsequent references to this battery will be abbreviated: MAP.

Problem of the Study

The specific problem of this study was to determine if the published norms for MAP are appropriate for native German students who take the test with a German translation of the tape recorded directions but with an untranslated answer sheet, as found in Appendix A. Because the adequacy of population norms is affected by the reliability of test scores, an ancillary problem was to estimate the reliability of MAP scores for the German students.

In the MAP standardization program, Gordon found that students who participated in chorus, band, and orchestra consistently made higher scores on MAP than did students who were non-participants.[21] As a result of this finding, separate norms were constructed by Gordon for students who are members of performing music groups or who have received private music instruction for at least one year (the term "musically select" was used to denote this type of student). Therefore, in the present study, the comparability of both total group norms and musically select group norms for German and American students was investigated.

---

21. Gordon, Musical Aptitude Profile, Manual, 65-67.

The Musical Aptitude Profile

MAP was selected for use in the present study because (1) German educators, as heretofore indicated, have expressed interest in the test, (2) the use of actual musical examples for the MAP test items complies with the principles of Gestalt psychology (which are highly regarded by educators in Germany), and (3) MAP has been thoroughly and carefully prepared, being one of the most important contributions to the continuing study of musical aptitude.[22] Thorndike refers to MAP as a "significant addition to the procedures available for assessing musical promise."[23] In reviewing MAP for the current edition of the Mental Measurements Yearbook, Lundin considered the test to be a valid instrument with a very adequate normative sample, and McLeish refered to MAP as "the best test of its kind on the market."[24] In the opinion of McLeish,

---

22. Paul R. Farnsworth, The Social Psychology of Music, 2nd ed. (Ames, Iowa: The Iowa State University Press, 1969), 205; and Paul R. Lehman, Tests and Measurements in Music (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968), 54.

23. Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education, 3rd ed. (New York: John Wiley and Sons, Inc., 1969), 374.

24. Oscar K. Buros, ed., The Seventh Mental Measurements Yearbook, 2 vols. (Highland Park, New Jersey: The Gryphon Press, 1972), I, 529.

MAP "conforms to all the criteria of excellence not only in test construction and validation, but in musicality [and] succeeds in measuring the high level functions of musical aptitude, and avoids the trap of measuring the effects of training."[25]

A complete description of MAP and the procedures used in establishing norms from the standardization program are included in the test manual.[26] Briefly, MAP comprises three main divisions or Total Tests: Tonal Imagery, Rhythm Imagery, and Musical Sensitivity. Tonal Imagery includes two subtests, Melody and Harmony, in which the subject is to decide whether a certain aspect of an "answer" phrase is "like" or "different" from a "question" phrase. In the Melody subtest, and in all but two of the remaining subtests, the musical phrases are performed on a violin; the exceptions are the Harmony subtest and the Phrasing subtest (in the Musical Sensitivity division) in which a violin and 'cello are heard together. In the Rhythm Imagery division, there are also two subtests: Tempo and Meter. For these subtests the subject is to decide whether a certain aspect of an "answer" phrase is the "same" as or "different" from a "question" phrase.

---

25. Ibid., 529-530.

26. Gordon, Musical . . ., Manual, 3-23.

In the Musical Sensitivity division, the subject is to make a judgment regarding the better of two short musical selections. Three subtests are included: Phrasing, Balance, and Style. In the Tonal Imagery and Rhythm Imagery tests, the subject listens to the test items and indicates an objective response, and in the Musical Sensitivity test, the subject indicates a preferential response. The responses are indicated by filling in an oval on the answer sheet.[27] For each test item, an "in doubt" response oval is provided on the answer sheet. This provision helps to preclude guessing and significantly increases the reliability of test scores.[28] No knowledge of music reading is necessary to take the MAP battery.

A unique feature of MAP is that it was subjected to a three-year predictive validity study.[29] Based on this study, using (1) judges' evaluation of students' tape-recorded musical performance, (2) musical achievement test scores, and (3) teacher ratings as validity criteria, the zero order correlation between the three criteria

---

27. A MAP answer sheet is included in Appendix A.

28. Gordon, Musical . . ., Manual, 62-63.

29. Gordon, A Three-Year Longitudinal Predictive Validity Study of the Musical Aptitude Profile; Studies in the Psychology of Music, V (Iowa City: University of Iowa Press, 1967). The writer was privileged to serve in an integral role in the pursuance of this study.