

## INFORMATION TO USERS

**This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.**

**The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.**

**In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.**

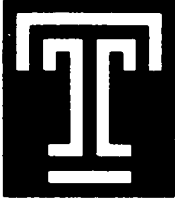
**Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.**

**Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.**

# UMI

A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
313/761-4700 800/521-0600





TEMPLE UNIVERSITY GRADUATE BOARD

Title of Dissertation:

A VALIDITY STUDY OF  
THE ADVANCED MEASURES OF MUSIC AUDIATION  
AMONG UNDERGRADUATE COLLEGE MUSIC MAJORS

Author:

Richard T. McCrystal

Read and Approved by:

*Maina Wright*  
*Stephen Walker*  
*Ray R. Meyer*  
*Noam Elkann*  
*Eileen Lynch*

Date submitted to Graduate Board: .....12/5/94.....

Accepted by the Graduate Board of Temple University in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Date *12/15/94* .....  
*[Signature]*  
(Dean of Graduate School)



A VALIDITY STUDY OF  
THE ADVANCED MEASURES OF MUSIC AUDIATION  
AMONG UNDERGRADUATE COLLEGE MUSIC MAJORS

---

A Dissertation  
Submitted to  
the Temple University Graduate Board

---

in Partial Fulfillment  
of the Requirements for the Degree  
DOCTOR OF PHILOSOPHY

---

by  
Richard T. McCrystal  
January, 1995

**UMI Number: 9527513**

---

**UMI Microform 9527513  
Copyright 1995, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized  
copying under Title 17, United States Code.**

---

**UMI**

**300 North Zeeb Road  
Ann Arbor, MI 48103**

ABSTRACT

A VALIDITY STUDY OF  
THE ADVANCED MEASURES OF MUSIC AUDIATION  
AMONG UNDERGRADUATE COLLEGE MUSIC MAJORS

by Richard T. McCrystal  
Doctor of Philosophy  
Temple University, 1995

Major Advisor: Dr. Edwin E. Gordon

The purpose of this study was to investigate the validity of the Advanced Measures of Music Audiation (AMMA), a test of music aptitudes for college undergraduate music majors. The three problems of the study were to establish the reliability of AMMA; to establish the immunity of AMMA to instruction, practice, and maturation; and to determine if AMMA scores could serve as predictors of success in music study.

The investigator administered AMMA twice to the same forty college undergraduate music majors. The first administration was at the beginning of the academic year and the second administration was at the end of the academic year. The results of the first administration were used to

determine predictive validity. The students' grade point average in applied music, aural skills, and written harmony were the validity criteria. The split-halves method was used to determine the reliability of the test. Both administrations of AMMA served as a pretest and posttest to determine the test's immunity to instruction, practice, and maturation.

The split-halves reliability coefficients were .82 for the pretest and .84 for the posttest. The gain in mean scores between the pretest and the posttest was not significant at the .05 level of confidence. The correlation coefficients between the AMMA pretest and each of the validity criteria were strongest when limited to only the highest and lowest AMMA scores. The highest and lowest AMMA scores were also better predictors of success in music than achievement in any of the validity criteria.

It may be concluded that AMMA is a reliable test. AMMA was also immune to instruction, practice, and maturation over the length of the study. It may also be concluded that the AMMA predictive validity coefficients found for high and low scoring students were more robust than were any of the intercorrelation coefficients found among the three validity criteria.

## ACKNOWLEDGEMENTS

The writer would like to express his gratitude to Dr. Edwin E. Gordon for his guidance and encouragement as chair of the dissertation committee. In addition, the writer wishes to express his thanks to Dr. Roger Dean, Dr. Eve Meyer, Dr. Stephen Willier, and Dr. Maurice Wright for their cordial willingness to share their knowledge and expertise to help make this study worthy.

The writer also wishes to thank Mr. Robert Parrish, Chair of the Department of Music, Trenton State College, Trenton, New Jersey. Both he and members of Trenton's theory department played a vital role in making this study possible. Also to be thanked are the many undergraduate students whose willing participation was central to the study.

My wife Margaret and our children Walter, Hilary, Margery, and Tobias were all supportive and happy for me to become a student again. I hope they will also never stop learning.

TABLE OF CONTENTS

	Page
ABSTRACT . . . . .	iii
ACKNOWLEDGEMENTS . . . . .	v
LIST OF TABLES . . . . .	viii

CHAPTER

1. INTRODUCTION . . . . .	1
2. RELATED RESEARCH . . . . .	12
Introduction . . . . .	12
A Longitudinal Predictive Study of the Advanced Measures of Music Audiation by Edwin E. Gordon . . . . .	13
An Investigation of the Validity of the Advanced Measures of Music Audiation With Junior and Senior High School Students, by David L. Fullen . . . . .	15
Effects of Training and Practice on Advanced Measures of Music Audiation Scores, by Steven G. Estrella . . . . .	19
3. DESIGN AND PROCEDURES . . . . .	27
Sample . . . . .	27
Design and Procedures . . . . .	28
Data Analysis . . . . .	29

4. RESULTS AND INTERPRETATION . . . . .	32
Results . . . . .	32
Interpretation . . . . .	36
5. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS . . . . .	42
Summary . . . . .	42
Conclusions . . . . .	44
Recommendations . . . . .	44
REFERENCE LIST . . . . .	46

## LIST OF TABLES

TABLE	Page
1. Split Halves Reliability Coefficients of the Students' Total Raw Scores and Split Halves Reliability Coefficients of the National Sample's Total Raw Scores . . . . .	33
2. Students' Pretest and Posttest Means and Standard Deviations for AMMA Total Raw Scores and the Mean and Standard Deviations for College and University Music Majors Nationally . . . .	33
3. Difference Method $t$ Test for AMMA Pretest-Posttest Total Raw Scores . . . . .	34
4. Entire Sample Correlation Coefficients of AMMA Total Mean Raw Scores with the Mean GPA in Applied Music, Aural Skills, and Written Harmony . . .	35
5. Group A, Group B, and Entire Sample Correlation Coefficients of AMMA Total Mean Raw Scores with the Mean GPA in Applied Music, Aural Skills, and Written Harmony . . . . .	35
6. Group A and Group B Means, Standard Deviations, and $t$ Values for AMMA Pretest Total Raw Scores and GPA in Applied Music . . . . .	36
7. Group A and Group B Means, Standard Deviations, and $t$ Values for AMMA Pretest Total Raw Scores and GPA in Aural Skills and Written Harmony . . . . .	36
8. GPA Correlation Coefficients for Applied Music, Aural Skills, and Written Harmony for the Entire Sample . . . . .	40

## CHAPTER 1

## INTRODUCTION

To warrant the assertion that the Advanced Measures of Music Audiation (AMMA) by Edwin E. Gordon (Gordon, 1989) is a test for stabilized music aptitudes would require establishing the validity of AMMA as a test of music aptitudes, particularly stabilized music aptitudes (Gordon, 1987, pp. 17-52). First, some preliminary questions need to be asked. Do aptitudes exist? If they exist, can they be identified and measured? Are there specific aptitudes for music? If there are specific aptitudes for music, how may they be described? Once described, how may they be further studied and understood? Presuming there are specific music aptitudes, does AMMA meet the criteria for a valid test of music aptitudes?

Do aptitudes exist? In both World War I and World War II, the United States found it had to rapidly change from a nation with a small peace time military establishment to a nation with an enormous military force. A major factor in successfully accomplishing this task, within a relatively brief time frame, was aptitude testing. Aptitudes, basic to military needs, were identified. Aptitude tests for those

aptitudes were developed. The resultant aptitude tests were administered to a large sample from the civilian population. Many from that sample were selected, based upon their aptitude test scores, for brief but intensive training in skill areas in which they scored high. The criterion validity of those efforts may be found in the success of the military establishment. From such successful results one may presume aptitudes exist, and that aptitudes can be identified, measured, and evaluated with a high degree of veridicality.

Are there music aptitudes? As with all human behaviors, one may safely presume there are aptitudes for music. The safety of this presumption is manifested in the many efforts, some quite successful, to identify and measure behaviors related to aptitudes for music (Seashore, 1919 & 1967; Seashore, Lewis & Saitveit, 1957; Wing, 1961; Shuter-Dyson & Gabriel, 1981; and Gordon, 1987). How many music aptitudes be described? Gordon wrote, "At the core of the nature of music aptitude is what traditionally has been referred to as the 'nature-nurture' issue" (Gordon, 1987, p.5). At issue is which is the greater factor regarding how one's music aptitude develops, nature (genetics) or nurture (environment). Gordon concluded the roles of the innate and environment regarding music aptitudes are not clear, and not known (Gordon, 1987, p. 7).

May we be certain aptitudes are innate? Two ways to address this question would be from either the biological view or the psychological perspective. A possible third approach would be at the juncture between biology and psychology. For example, studies have been conducted on song birds. Staddon, an animal behaviorist, noted that some bird species (e.g., song sparrows, *Melospiza melodia*) are little affected by early deprivation of song role models because their song information is clearly innate. On the other hand, some species (e.g., white-crowned sparrows, *Zonotrichia leucophrys*) develop severely abnormal songs if deprived of the opportunity to hear their specie's song during the first six months of life (Staddon, 1983, p. 398). Studies of other species, including humans, involving different sensory pathways, reported similar findings (Edelman, 1987, p. 302). That is, deprivation at critical growth periods results in abnormalities. Arms, a biologist, would say that animals are genetically (innately) predisposed to learn the behavior patterns of their own species (Arms, 1979, p. 616). Gordon, whose interest is at the interface between music and psychology, would add, a music environment for human infants that either encourages or discourages music behavior, will have a positive or negative influence on the infant's music behavior (Gordon, 1987, p. 8).

Longitudinal studies conducted on human monozygotic (same egg) twins, separated at birth and reared apart,

provides support for the innate claim for human characteristics as reported by Bouchard, Lykken, McGue, Segal, and Tellegen (1990). According to their study, among such twins about 70% of the variance in IQ was found to be associated with genetic variation. What may be somewhat surprising are the findings reported in the study on other traits of monozygotic twins reared apart. On multiple measures of personality and temperament, occupational and leisure-time interests, and social attitudes, twins reared apart were about as similar as monozygotic twins reared together. It is a strong argument for the power of the genetic or innate factor. However, it is not meant to demean the power of the environmental factor. Both, as Gordon would assert, are important.

There is a third factor to be considered in the nature-nurture equation, a question that seems to have a potent influence; when, in chronological time, is the organism genetically most open to be influenced by its environment? Primitive species may be less affected in their development by deprivation at critical chronological periods, but a higher life form is affected. The white-crowned sparrow is one example of how deprivation of song role models can affect its song production at a critical time in its development. For humans, Gordon has reported on test results that indicate music aptitude stabilizes around the age of 9. Between birth and age 9 music aptitude is developmental, or more readily

influenced by environment. In addition, the earlier children have favorable experiences with music, the more susceptible their music aptitude is to change. It may even approach its innate potential (Gordon, 1987, p. 9). The identification of developmental and stabilized music aptitudes strongly suggests the need for a variety of tests to be designed for either developmental or stabilized music aptitudes.

Related to the nature-nurture issue is a learning loop, and its place in time. Staddon wrote of such a learning loop in which birds develop their species' song. Initially this loop in birds is a genetically influenced template for a specific type of song. But a template is not enough. The other factor in the loop is practice (Staddon's underline). In order for practice to be effective it must take place during critical growth periods. There is a striking similarity between Staddon's learning loop for bird song and Gordon's description of a "babble" stage in the early development of music aptitudes. For birds, Staddon wrote, "The final song is evidently built up by a circular process in which the bird sings, hears the result of its singing, and slowly modifies what it sings until its production matches the stored template" (Staddon, 1983, p. 400). Staddon thought this was analogous to a musician programmed to learn a single Baroque piece. He would be immune to the "charms of Brahms or the Beatles, but would fixate instantly on anything by Bach or Vivaldi" (Staddon, 1983, p. 400).

According to Staddon, "This loop can be interrupted in two places: either by disrupting the motor apparatus, to prevent practice, or by deafening the bird, to prevent its hearing its own production" (Staddon, 1983, p. 400). An experimentally deafened bird will develop a normal subsong (from its template), but it will never evolve it into a normal adult song. This experiment offers strong support for the importance of auditory feed-back, and is similar to what Gordon has concluded from his investigations. Gordon reported how critical early informal and formal experiences in music are to maintain a child's level of music aptitude. He wrote, "What is known is that regardless of the level of music aptitude a child is born with, he must have favorable early informal and formal experiences in music in order to maintain that level of potential" (Gordon, 1987, p. 8). He added that one cannot increase his innate level of music aptitude with favorable early experiences in music. A lack of early experiences, however, could adversely affect his music aptitude, whether potentially high or low, to perhaps only slightly more than nothing.

Another dimension regarding babble was reported in a recent study on hearing infants between 7 and 10 months of age and severely deaf infants of similar age (Petitto & Maretette, 1991, pp. 1493-1496). While hearing infants engaged in spontaneous speech babble, the deaf infants engaged in spontaneous manual (sign language) babble. This

study not only supports the concept of babble, but it indicates there is an underlying innate capacity for language in the human species. If the customary speech modality for language babble is adversely affected, the mode in which language manifests itself seems optional. Prudence would suggest there is also an underlying innate capacity for music in humans.

Any study about music aptitude must address the meaning of audiation, a term put forward by Gordon to fill a serious gap in our language (Gordon, 1976, p. 6). It is to music what auding is to a spoken language; that is, auding is what we do when we hear and process the sound of language inside our heads without those sounds being physically present on the outside (Taylor 1964, p. 4). When sounds are physically present and heard one is also perceiving. However, perceived sounds need to be processed through auding or audiation to be intelligible.

We audiate when we listen to music for its syntactical tonal and rhythm patterns.

We audiate when we read notation and hear what we read, with or without making an external sound.

We audiate when we write music from dictation.

We audiate when we recall music from our past.

We audiate when we write music we have recalled.

We audiate when we improvise music on the spot.

And we audiate when we write down the music we have created or improvised (Gordon, 1988, pp. 10-13).

The powerful role audiation plays in any processing of sound as music, as exemplified above, leads to the conclusion that its role in how one describes and measures music aptitudes is formidable. It supports Gordon for his choice of the word "audiation" in AMMA rather than "aptitudes."

The distinction between developmental and stabilized music aptitudes is critical to this study, for stabilized music aptitudes are believed to be more immune to practice than developmental aptitudes. Practice customarily means to take the time necessary to hone one's performing and aural skills, mostly through the exercise of those skills. To qualify as a test of stabilized music aptitudes, AMMA scores should not be significantly improved by practice, instruction, or maturation.

Another factor critical to this study is the distinction between aptitude and achievement. Aptitude is potential, or how suited one is to a purpose. Achievement is what one accomplishes. They are not the same, for achievement is dependent upon aptitude. It is not uncommon to hear it said that aptitude may be measured by achievement. The strength of that argument is, one cannot achieve highly without a correspondingly high aptitude for what it is one achieved so well. The weakness of that argument is, it ignores the issue of persons with high aptitude and low achievement. According to Gordon, in a typical school there are many students (perhaps as many as 50%) with high music aptitude who are not

receiving any formal music instruction. Conversely, there are also many students with low music aptitude who are receiving formal music instruction (Gordon, 1988, p. 6). If those responsible for mobilizing the United States in 1917 and 1941 had relied upon achievement to identify persons with the needed aptitudes for defense, it is likely the screening of the civilian population for prospective candidates for training would have been severely limited. The results could have been catastrophic.

Katherine Mansfield, the British short story writer, wrote in her Journal, "I want to be all that I am capable of becoming." (Bartlett, 1980, p. 810). She suggests the question, how will one know what it is one is capable of becoming without knowing one's aptitudes. Measurements taken with AMMA may help persons to evaluate what they are capable of attaining in music.

This study will be conducted to determine the predictive validity of AMMA by how well AMMA scores relate to achievement in music. Student's Grade Point Average (GPA) in applied music, aural skills, and written harmony will be compared with subject's Total AMMA scores. If there is a strong positive relationship between GPA and AMMA, it could mean the subject's achievement in music could be predicted by how well he scored on AMMA.

A factor that separates aptitude from achievement is practice. Music achievement can be influenced by practice,

up to a point. Stabilized music aptitudes are thought to be immune to practice, unlike developmental music aptitudes (Gordon, 1987, pp. 40-52). This study will investigate whether or not AMMA scores are influenced by practice.

Research conducted on the Music Aptitude Profile (MAP), which is a test for stabilized music aptitudes, indicates there are small increments, related to chronological age, in raw scores. However, a person's percentile rank retains its relative position even though his raw scores may show some increase related to an increase in age. Consequently, different norms are reported in MAP according to chronological age (Gordon, 1965 pp. 90-113).

The norms established for AMMA are the same for undergraduate and graduate students, although norms differ between music majors and non-music majors (Gordon, 1989, pp. 28-31). Apparently, chronological age differences at the college level have no effect on AMMA raw scores. The increase in raw scores on MAP could be because the population MAP was designed for experiences a final growth spurt in brain development around the age of 15 (Epstein, 1978, pp. 343-370). The age group for whom AMMA was designed is past having brain growth spurts. The biology of the brain may help to explain why the norms for AMMA are the same for undergraduate and graduate students.

Problems

The following are the specific problems of the study.

- 1) To establish the reliability of AMMA.
- 2) To establish the level of immunity of AMMA to practice, instruction, and maturity in applied music, aural skills, and written harmony.
- 3) To determine if AMMA scores could serve as a predictor of success in music study.

## CHAPTER 2

## RELATED STUDIES

## Introduction

Related studies of Gordon's 1989 Advanced Measures of Music Audiation (AMMA) are few in number, because it is a relatively new test, and the only advanced test of music aptitudes currently in print. Initially, AMMA was designed and normed for college and university undergraduate and graduate students. Subsequently, AMMA norms have been established for senior high school and for 7th and 8th grade levels, thus increasing its potential use. The 20 minutes it takes to administer AMMA, and its ease of scoring makes AMMA an attractive measuring instrument to use with students who range in age from 12 to 20 and beyond.

Perhaps the strongest reason AMMA has suffered from critical neglect is because it is specifically intended to measure audiation, the basis of music aptitudes. The concept of audiation is apparently not widely understood, and consequently, it cannot be widely appreciated.

A Longitudinal Predictive Validity Study of the  
Advanced Measures of Music Audition

Edwin E. Gordon's predictive validity study of AMMA (Gordon, 1990) was among the first published studies of AMMA. Initially, the study involved 225 undergraduate and graduate students who were members of the orchestra, concert choir, and band of the Esther Boyer College of Music of Temple University. They were given AMMA at the beginning of the academic year. Of the 225 students, 114 completed the study by recording a short etude at the end of the academic year. The etude was composed especially for the study and was keyed for a variety of instruments and voices. The students had approximately 10 days to prepare the etude without teacher help. How well it was performed served as the study's validity criterion.

The students' recorded etude was evaluated independently by a panel of 3 judges. Each judge used a rating scale for each of three dimensions to be evaluated. The three dimensions were Tonal or Instrumental Technique, Rhythm, and Expression. The Tonal part of the first dimension (based on intonation) was for all instruments except piano. Pianists were rated on the Instrumental Technique part of the first dimension.

A comparison of the judges' mean ratings of the recorded etude with the theoretical mean for each dimension, and the

theoretical mean for the combined dimensions indicated that the judges were able to discriminate and agree on the performances of the etude rather well (Gordon, 1990, p. 8). In a scale of from 1 up to 5 that was used for each dimension, the theoretical mean for each dimension was 3.0.

The ratings of all the judges combined for all the dimensions combined had a reliability coefficient of .84. In addition, the intercorrelations of the combined judges' ratings of the students' performances of the recorded etude ranged from .88 to .89, which is substantial (Gordon, 1990, p. 10). Such high reliability intercorrelations were important because the scores from the rating scales were the validity criterion for the study.

The result of the study was that the longitudinal predictive validity coefficient between the Total scores on AMMA with the judges' combined ratings of the combined dimensions of the rating scale was .82 (Gordon, 1990, p. 11). This indicates that 67% (the square of .82) of the reason or reasons for the students' success in music performance can be predicted by the students' Total test scores on AMMA.

A further indication of the discriminative predictive power of AMMA was arrived at by comparing the relationship between the highest scoring students and the lowest scoring students on the Total AMMA test results. When the high scoring and the low scoring group test results were compared with the combined judges' ratings of the etude performance it

corroborated that students who scored high on AMMA performed the etude significantly better than those who scored low on AMMA (Gordon, 1990, p.12).

The AMMA longitudinal predictive validity coefficient Gordon obtained in his study (.82) is similar to the longitudinal predictive validity coefficient of .81 he reported when the Music Aptitude Profile (MAP) results were combined with the Instrument Timbre Preference Test (ITPT) results (Gordon, 1984, p. 48). That is, AMMA by itself was able to deliver a predictive validity coefficient that took the combined results of MAP and ITPT to equal. To give this some perspective, MAP requires 150 minutes to administer and ITPT requires 30 minutes to administer. AMMA requires 20 minutes to administer or about 11% of the time needed to administer both MAP and ITPT. All three tests are designed to be given to groups of students.

An Investigation of the Validity of the  
Advanced Measures of Music Audiation With  
Junior and Senior High School Students

David L. Fullen's study (Fullen, 1993) is of interest both as a validity study of AMMA and as a study concerned with secondary school students for whom norms have been established. He investigated two problems. The first was to determine the predictive validity of AMMA for secondary

students. The second was to determine the effect participation in choral ensembles had on the students' AMMA scores.

Fullen's study involved 126 students in grades 7 and 8 (junior high school) and 91 students in grades 9 through 12 (senior high school), all of whom were in a choir. The 217 students were first given AMMA at the beginning of the academic year. Ten days later they were all given AMMA for a second time to establish retest reliability coefficients. AMMA was administered a third time at the end of the first semester to 104 students in four of the five junior high schools who had choir for only one semester. The remaining 22 junior high students in the fifth junior high school and all 91 senior high school students took the third administration of AMMA at the end of their second semester (Fullen, p. 24).

The criterion measure for the study was the independent rating by two judges of recordings of two melodies sung by the students. These recordings were made about the same time the students were given AMMA for the third time. Fullen composed four melodies for the study, two for junior high students and two for senior high students. He also constructed the rating scale that was used. Three dimensions of the students' vocal performance were rated: 1) tonal, 2) rhythm, and 3) expression.

The first two administrations of AMMA that were 10 days apart established the retest reliability coefficients for both groups. The retest Total AMMA scores reliability coefficient was .81 for the senior high school students (Fullen, p. 29). It compares well with the retest Total AMMA scores reliability coefficient of .83 for undergraduate and graduate non-music majors (Gordon, 1989, p. 41). Gordon reported split halves reliabilities for senior high school students' Total AMMA scores as .84 (Gordon, 1989, p. 40). That, too, is quite close to the .81 retest reliabilities Fullen reported.

The retest Total AMMA scores reliability coefficient for junior high school students was .72. Gordon reported a split-halves reliability coefficient of .85 for this cohort (Gordon, 1989, Manual Addendum). Although the obtained junior high school reliability coefficient of .72 is lower than the obtained senior high school reliability coefficient of .81, it is still considered adequate.

The rating scales developed by Fullen were evaluated by the interjudge reliability coefficients as an estimate of the reliability of the judges' ratings. The interjudge reliabilities for the rating scale were high, .94 for junior high school students and .95 for senior high school students.

The correlations between the Total AMMA raw scores from the first test and the combined judges' ratings of all dimensions combined were low. The longitudinal predictive

validity of junior high students was .25, and for senior high students .24. In trying to determine the reason for the low predictive validity, Fullen speculates that the content of the rating scale may have been inappropriate, that the judges were too lenient in rating student performance, or that both the rating scale and the judges were at fault (Fullen, p. 43).

In contrast, the means for the combined judges' ratings on each dimension and all dimensions combined of junior high school students who scored in the top 20% on AMMA were significantly higher than the means of students who scored in the lower 20%. Senior high school students who scored in the upper 20% on AMMA were rated significantly higher than the lower 20% on the tonal dimension (Fullen, p. 45).

To investigate the effects of music instruction on AMMA scores, Fullen correlated the scores of the first administration of AMMA with scores from the third administration of AMMA. The results indicate that the students generally maintained their relative position on the test through the period of instruction, whether it was for one semester or two semesters (Fullen, p. 42). The difference of the Total AMMA mean raw score gain between test administration one and test administration three was 1.15 for junior high school students and 1.62 for senior high school students. None of these differences was statistically significant (Fullen, p. 42).

It was concluded by Fullen that AMMA scores accurately predict music achievement for groups of secondary school students who demonstrate either high or low levels of music aptitude. He found AMMA scores predict less well for individual students. In addition, Fullen concluded that scores on AMMA are not sensitive to instruction, achievement or familiarity with taking the test itself (Fullen, p. 46).

#### Effects of Training and Practice on Advanced Measures of Music Audiation Scores

This study by Steven G. Estrella (Estrella, 1992) departs from the above studies in some significant ways. It was an experimental study, that is, it provided treatment to experimental groups, but not to control groups. The purpose of the study was to investigate the effects of training and practice in tonal audiation skills on music audiation test scores. The problem of the study was to determine if pretest-posttest gain scores for students who received treatment were significantly different from pretest-posttest gain scores for students who did not receive treatment. The Tonal subtest and the Rhythm subtest of AMMA were the locus for any differences in scores, and those scores were the two criterion measures used in the study.

There were two experimental groups and two control groups. One pairing of experimental and control groups was

There were two experimental groups and two control groups. One pairing of experimental and control groups was at the university level and the other pairing of experimental and control groups was at the senior high school level. All four groups were a mix of music majors and non-music majors. The university level groups were in Philadelphia and the senior high school groups were in British Columbia.

Thirty-seven undergraduate music and non-music majors and 39 high school seniors were administered AMMA at the beginning of the 1991 Spring term. From that sample, 31 undergraduates and 30 high school seniors took AMMA at the end of the 1991 Spring term. Random assignment to experimental and control groups was not possible and volunteerism was necessary to obtain control groups willing to be tested twice and experimental groups willing to be both tested twice and treated (Estrella, p. 17).

The two experimental groups received treatment on tonal audiation only. Tonal audiation training was done on Macintosh computers through digitized voice sounds for patterns using tonal syllables, and through synthesized clarinet sounds for patterns not using tonal syllables. The students were taught to identify tonal pattern functions and they were evaluated by how well they developed that skill (Estrella, p. 20). They were also taught to determine if two tonal patterns were the same or different. The tonalities used for the training and practice patterns were: major,

minor, aeolian, mixolydian, dorian, phrygian, and lydian, in that order (Estrella, p. 20). Both experimental groups received treatment on the seven tonalities in twenty-six, thirty minute sessions, a potential average of a little less than two hours per tonality.

For the purpose of preliminary analyses, two two-way factorial analyses of variance (ANOVA) were used to determine if there was a main effect for site (high school vs. university), a main effect for assignment (experimental vs. control), and an interaction effect between site and assignment at the .05 level of confidence separately for both the AMMA Tonal and AMMA Rhythm pretest scores. The results of the preliminary analyses of the pretest AMMA Tonal scores showed no main effect for site, no main effect for assignment, and no interaction effect between site and assignment.

Results of the preliminary analyses of the pretest AMMA Rhythm scores also showed no main effects for site or assignment and no interaction effect as well. Estrella concluded from this that at the beginning of the study no significant differences in AMMA Tonal or Rhythm scores existed among the four groups (Estrella, p. 28).

For the principal analyses, two two-way ANOVA were used to determine if there was a main effect for site, a main effect for treatment, and an interaction effect between site and treatment at the .05 level of confidence for either or

both the AMMA Tonal and AMMA Rhythm pretest-posttest gain scores. The results of the principal analyses of the AMMA Tonal pretest-posttest scores showed only a main effect for treatment. Estrella wrote that because there was no significant interaction between site and treatment for AMMA Tonal pretest-posttest gain scores, the main effect for treatment becomes relevant (Estrella, p. 29). That is, the results appear to Estrella to indicate that members of both the experimental groups improved their tonal scores significantly more than did members of both the control groups. It suggests that the two experimental groups profited from the intervention. Estrella added, "The main effect, however, cannot be interpreted as evidence of the effectiveness of the intervention because of the loss experienced by [the university control group] and the gain experienced by [the high school control group]" (Estrella, p. 30). He concluded, "The most confounding aspect of the results is the dramatic pretest-posttest loss for the university control group [-1.45] coupled with the modest pretest-posttest gain for the high school control group gained [.66]" (Estrella, p. 32).

Results of the principal analyses of the AMMA Rhythm pretest-posttest gain scores showed a main effect for treatment plus an interaction effect between site and treatment. Estrella said, "Because there was a significant interaction between site and treatment for AMMA Rhythm

posttest-pretest difference scores, the main effect for treatment becomes less relevant" (Estrella, p. 38).

Estrella thought the principal analyses evaluated only the difference between the differences of the separate groups, and supplementary analyses were needed. He used four two-tailed  $t$  tests of mean differences to determine if score changes experienced by the four groups (two experimental and two control) were statistically significant at the .05 level of confidence. The results indicate that the decrease of AMMA Tonal scores was significant for the university control group and the increase significant for the high school experimental group. Score changes for the university experimental group and high school control group were not significant (Estrella, p. 32).

Estrella attributes the divergent performances of the two experimental groups to two extreme difference scores in the university experimental group. Those two extreme difference scores inflated the standard deviation and depressed the paired  $t$  value. Estrella surmised that the small sample size ( $N=11$ ) of the university experimental group was also "a reason why the mean difference was not even close to significant" (Estrella p. 34).

The pretest-posttest difference scores on the AMMA Rhythm test were complicated, according to Estrella, because the two control groups experienced score changes in opposite directions. That is, the university control group scores had

a negative gain of  $-.85$  and the high school control group scores had a positive gain of  $.73$  (Estrella, p. 38). He believed further analysis was needed to determine the significance of score changes experienced by the four groups. Again he used four two-tailed  $t$  tests of mean differences. The results indicated that the gain for the university experimental group was significant but the score changes for the remaining three groups, according to Estrella, were attributable to chance.

Estrella questioned why the university experimental group had significant differences on AMMA Rhythm pretest-posttest gain scores and not on AMMA Tonal scores because their treatment was with tonal audiation skill training and not rhythm audiation skill training. He thought the high intercorrelations between the AMMA subtests was one possible reason for the increase in rhythm scores but not tonal scores (Estrella, p.40). Estrella went on to ask why the high school experimental group did not also have significantly higher rhythm scores. He concluded that the divergent performance of the two experimental groups on the rhythm test scores was probably due to site related factors. For example, the university students received more instruction in music related classes than the high school students.

The conclusions drawn by Estrella were different from the conclusions found in the other two related studies. For example, the negative gain for the control groups inflates

the positive gain for the experimental groups. Nevertheless, Estrella concluded AMMA tonal and rhythm scores can be improved by training in tonal audiation. The investigators of the other two related studies found AMMA scores did not improve significantly through training. Estrella did say that the use of control groups was probably unnecessary because maturation is not a factor regarding AMMA scores (Estrella, p. 50).

Estrella concluded, "1. Tonal audiation ability, as measured by AMMA, can be improved through specialized training in tonal audiation. 2. Rhythm audiation ability, as measured by AMMA, can be improved through a combination of specialized training in tonal audiation and participation in college-level music training" (Estrella, p. 49). He recommends the use of AMMA for predicting success in music performance. Estrella, however, cautions against using AMMA as a diagnostic measure for specific dimensions of music learning (tonal and rhythm).

In addition, Estrella recommends that a measuring instrument be developed to measure separately "specific audiation abilities of college students" (melodic and harmonic dimensions of tonal audiation, and tempo and meter dimensions of rhythm audiation) that would be modeled on Gordon's Music Aptitude Profile (Estrella, p. 49). Estrella said, "Previous research, using MAP [Music Aptitude Profile] and other audiation tests, has indicated that at least three

independent dimensions of music audiation exist" (Estrella, p. 51). He claims that the results of his study do not support the existence of multiple independent dimensions of music audiation. Instead, he believes his study indicates that audiation is unidimensional in persons aged 17 to 21. He did admit, "Such a provocative conclusion, however, would be premature because of the unique design and scoring procedures of AMMA as compared with more thoroughly researched audiation tests" (Estrella, p. 51).

..

## CHAPTER 3

## DESIGN AND PROCEDURES

## Sample

The sample in this study was drawn from an initial group of 65 undergraduate music majors attending a state college in New Jersey. From the initial cohort the investigator obtained 40 matched pretest and posttest AMMA scores. The protocols for use of human subjects were observed. Approximately 90% of the 40 subjects with matched test scores were Caucasian. In addition, most of the students were from New Jersey. Most received their pre-college music instruction in the local public schools. They were all enrolled as music education majors, and most were instrumentalists. The sample was a good representation of the music department's total student population. In addition, the obtained norms of the sample closely matched the national norms for college music majors.

According to the college undergraduate bulletin, all students admitted as music majors must have satisfactorily high math and verbal scores from either the Scholastic Aptitude Test (SAT) by the College Entrance Examination Board

or from the American College Testing Program (ACT). In addition, the requirement for admission to the department of music is by satisfactorily passing staff developed achievement tests in performance and aural skills. Standardized tests of music aptitudes are not used as part of the department's admissions policy.

### Design and Procedures

#### Design

First, it was necessary to determine whether or not AMMA is a reliable test of stabilized music aptitudes among a select group of college music majors. Only then could the next stage of the study be conducted, which was to determine whether or not AMMA is a valid test of stabilized music aptitudes among the college music majors tested, and whether or not AMMA could be used to predict performance in applied music, aural skills, and written harmony.

#### Procedures

AMMA was administered twice, to undergraduate students enrolled in aural skills classes and written harmony classes. The first administration was early in the fall term and the second administration was near the end of the spring term. The two test results served as the study's pretest and posttest.

The reliability of both the pretest and the posttest was determined by the split-halves method. To obtain the split-halves data a special scoring mask was used because of the unique way AMMA is scored, which is one of AMMA's virtues.

To determine the immunity of AMMA scores to practice and maturity a comparison was made between the Total mean raw scores of the pretest with the Total mean raw scores of the posttest. During the interval between the pretest and the posttest the students experienced most of an academic year's study of applied music, aural skills, and written harmony.

To investigate the predictive validity of AMMA, the pretest Total mean raw scores were correlated with the subject's mean GPA in applied music, aural skills, and written harmony. The strength of the obtained correlation coefficients would indicate how well AMMA can predict achievement in the three subject areas.

### Data Analysis

#### Reliability of Students' AMMA Scores

Correlation coefficients were calculated to estimate the reliability of the test results with the split-halves method on both the pretest and posttest Total scores. In addition, the pretest and posttest means and standard deviations were compared with the national sample's mean and standard deviation scores for closeness of fit.

### Effects of Instruction on Students' Posttest AMMA Scores

To determine if AMMA was affected by instruction in applied music, aural skills, and written harmony, a  $t$  test was used for the pretest-posttest Total raw scores to determine the significance of any gain scores.

### Predictive Validity

A correlation coefficient was calculated between the students' AMMA Total raw scores and their GPA in applied music, aural skills, and written harmony. Their GPA in those three subjects served as the validity criteria for the predictive validity of AMMA. To obtain the GPA in applied music, aural skills, and written harmony, the students' permanent records were accessed and their GPA averaged. Letter grades were converted to numerical grades according to the grading system published in the college's undergraduate bulletin. Those grades represented a one to three year period of study. The average GPA for each student was then correlated with each student's pretest AMMA Total raw score.

The resultant coefficients for the entire sample warranted a further investigation that led to examining the coefficients of the top 20% of the students (termed Group A) and the bottom 20% (termed Group B). In addition,  $t$  values were obtained to establish the significance of relationships between the two groups on their AMMA scores and GPAs.

The GPA coefficients obtained for the entire sample also served as an invitation to compare the correlation coefficients among the three subjects that served as the validity criteria.

## CHAPTER 4

## RESULTS AND INTERPRETATION

## Results

Reliability

The reliability of a test should be established before its validity can be determined. The reliability of the students' scores in this study may be viewed in two ways. One is by using the split-halves reliability coefficient method and the other is by matching the students' obtained mean and standard deviations with the mean and standard deviations of the national sample of undergraduate and graduate music majors.

The students' split-halves reliability coefficients are shown in Table 1. The obtained Total scores reliability coefficient for the pretest was .82 and .84 for the posttest. The split-halves reliability coefficient of the Total scores for the national sample was .88 (Gordon, 1989, p. 40). Reliability coefficients generally range between .00 and .95. It may be concluded that both test results are reliable and relatively free from the vicissitudes and distractions within and without the test room.

Table 1. Split Halves Reliability Coefficients of the Students' Total Raw Scores and Split Halves Reliability Coefficients of the National Sample's Total Raw Scores

	Students	National Sample
Pretest	.82	
Posttest	.84	.88

A comparison of the means and standard deviations of the scores obtained in this study and those of the national sample of undergraduate and graduate music majors (Gordon, 1989, p. 44) is shown in Table 2.

Table 2. Students' Pretest and Posttest Means and Standard Deviations for AMMA Total Raw Scores and the Mean and Standard Deviations for College and University Music Majors Nationally

	Pretest	Posttest	Nationally
Mean	58.2	60.0	59.1
S. D.	6.58	6.90	7.41

### Effects of Instruction

In comparing the students' pretest and posttest Total mean raw scores, the difference was found to be less than two points. As shown in Table 3, using the difference method  $t$  test, a  $t$  value of 1.93 was obtained. To be statistically significant at .05 confidence level, the  $t$  value would have to have been 2.02 or higher. Thus, the obtained difference of 1.93 was not statistically significant.

Table 3. Difference Method  $t$  Test for AMMA Pretest-Posttest Total Raw Scores

	Pretest	Posttest	
n	40	40	
Mean Raw Score	58.2	60.0	
Difference			-1.8
$t$ Value			1.93*

\*.05 >2.02

The lack of a statistically significant difference between the students' pretest and posttest Total mean raw scores suggests that the instruction received by the students during the fall term and the spring term had little or no influence on how they performed on AMMA. If AMMA were an achievement test one would expect a statistically significant difference between scores at the beginning and at the end of an academic year. This was not the case. If AMMA is a test of stabilized music aptitudes one would expect the test to be immune to practice and maturation. Statistically, this seems to be the case.

#### Validity

To investigate the predictive validity of AMMA, correlation coefficients were obtained between the Total AMMA scores and GPA in applied music, aural skills, and written harmony. As can be seen in Table 4, the coefficients between the AMMA Total scores and the three validity criteria were low.

Table 4. Entire Sample Correlation Coefficients of AMMA Total Mean Raw Scores with the Mean GPA in Applied Music, Aural Skills, and Written Harmony

	Applied Music	Aural Skills	Written Harmony
	.28	.04	.02

Because the coefficients were low and served no predictive purpose, the top 20% of the sample (called Group A) and the bottom 20% of the sample (called Group B) were examined. The AMMA Total raw scores of Group A and of Group B show a much higher relationship with the three validity criteria. They are compared in Table 5.

Table 5. Group A, Group B, and Entire Sample Correlation Coefficients of AMMA Total Mean Raw Scores with the Mean GPA in Applied Music, Aural Skills, and Written Harmony

	Applied Music	Aural Skills	Written Harmony
Group A	.56	.49	.72
Group B	.47	.09	.52
Entire Sample	.28	.04	.02

To further understand the performance of Group A and of Group B, the means, standard deviations, and  $t$  values for AMMA Total pretest scores and the GPA for each of the three validity criteria were obtained. They are shown in Table 6 for applied music and in Table 7 for aural skills and written harmony.

Table 6. Group A and Group B Means, Standard Deviations, and  $t$  Values for AMMA Pretest Total Raw Scores and GPA in Applied Music

Group	AMMA		Applied Music	
	A	B	A	B
Mean	66.8	48.4	3.6	3.4
S. D.	1.83	3.78	0.41	0.72
$t$ Value	7.62*		2.16*	

\*.05 >1.90

Table 7. Group A and Group B Means, Standard Deviations, and  $t$  Values for AMMA Pretest Total Raw Scores and GPA in Aural Skills and Written Harmony

Group	AMMA		Aural Skills		Written Harmony	
	A	B	A	B	A	B
Mean	66.8	48.4	2.7	2.8	2.7	2.9
S. D.	1.83	3.78	0.85	0.82	1.15	0.68
$t$ Value	7.62*		-0.31		-0.74	

\*.05 >1.90

The  $t$  values, of course, were statistically significant for the AMMA scores, and they were also statistically significant for the GPA in applied music. The  $t$  values for the GPA in aural skills and written harmony were not statistically significant.

### Interpretation

#### Sample

The data appears to support the assumption that the participating students were representative of the population

from which they were drawn. The participants numbered more than one-third of all those enrolled as undergraduate music majors. The AMMA scores generated by the sample also had a normal distribution.

### Reliability

The split-halves reliability coefficients obtained for the Total AMMA raw scores were .82 for the pretest and .84 for the posttest. They compare favorably with the split-halves reliability coefficient of .88 for the national sample (Gordon 1989, p. 40). Gordon also reported a split-halves reliability coefficient of .88 in his related study (Gordon 1990, p. 7). Fullen obtained a Total AMMA raw scores test-retest reliability coefficient of .72 for junior high school students, and .81 for senior high school students (Fullen, p. 29). Estrella did not report any test reliability coefficients. Correlation coefficients between .70 and .90 are considered high; that is, they indicate a marked relationship between any two factors. We may conclude from the reliability coefficients reported above that AMMA is a reliable test.

### Effects of Practice

The failure to obtain statistically significant different mean scores between the pretest and posttest supports the assertion that AMMA is a test of stabilized

music aptitudes rather than a test of music achievement. AMMA appears to be immune to the effects of practice, study, and maturation. The differences between the mean scores of the pretest and posttest must be attributed to chance.

### Validity

The somewhat low predictive validity for the entire sample between the AMMA Total scores and the three validity criteria may be traced to the different powers of discrimination involved. For example, it states in the undergraduate bulletin that the grade for "average" is "C" or 2.0. The grading system used by the college is A to F and it is suitable for grading a normal distribution. The percentage of students, however, averaging at or above the grade of "B" (above average) was 98% for applied music, 90% for aural skills, and 98% for written harmony. In a normal distribution, such as the one that was obtained from the entire sample's AMMA Total scores, about 68% of the sample may be considered as average and approximately 16% as above average. The GPAs indicate that 90% to 98% of the entire sample were above average. That leaves 2% to 10% who were average or below average. Such a lack of variability within each of the three validity criteria in contrast with the excellent variability obtained from the AMMA Total scores prevents any substantial correlation between AMMA Total scores and the entire sample's three validity criteria.

Group B's low correlation coefficient between AMMA Total scores and aural skills GPA may be explained by looking at two scores in the Group. Group B's lowest AMMA Total score did not have a correspondingly low aural skills GPA, and Group B's highest AMMA Total score did not have a correspondingly high aural skills GPA. Consequently, it distorted the coefficient for the group .

Another factor that resulted in low correlation coefficients for the entire sample's AMMA Total scores and the three validity criteria was that AMMA tests audiation skills while the three validity criteria may not be graded according to how well one audiates. The students' general intelligence seems to play a greater role than their audiation regarding what grade they will receive in applied music, aural skills, and written harmony. One of Gordon's major contributions to the psychology of music is his finding that music aptitudes and general intelligence correlate low. Consequently, music aptitudes cannot be accurately predicted by performance on a general intelligence test or on an academic achievement test. Nor can general intelligence be accurately predicted by performance on a music aptitudes test (Gordon, 1987, p. 117). Logically, scores on an audiation test will not correlate highly with grades based upon general intelligence associated behaviors.

The grade one receives in applied music cannot be used with much confidence to predict grades in either aural skills

or written harmony. Their correlation coefficients are shown in Table 8. They were lower than one might expect for such highly related kinds of activities, especially aural skills and applied music. Moreover, the coefficients that describe the relationship between applied music, aural skills, and written harmony are considerably lower than the coefficients that describe the relationship of music aptitudes to each of the validity criteria.

Table 8. GPA Correlation Coefficients between Applied Music Aural Skills, and Written Harmony for the Entire Sample

	Applied Music	Aural Skills	Written Harmony
Applied Music	--	.40	.25
Aural Skills	.40	--	.28
Theory	.25	.28	--

Interestingly, the low coefficients shown in Table 8 are similar to the low coefficients between music aptitudes and general intelligence. Those coefficients are .31 and .34 (Gordon, 1988, p. 80). Gordon had also reported low coefficients between music aptitudes, arithmetic, and mathematics skills of between .20 and .40. In contrast, Gordon wrote, "It is well known that the correlation between scores on general intelligence tests and academic achievement tests are approximately .90 (Gordon, 1988, p. 85). It brings into question the way rhythm is typically taught, that is as note value counting or as an arithmetical skill. Arithmetic

is useful to show ratios among durations, however, musically rhythm is audiated as patterns in a context or a syntax, but not as isolated durations.

Typically, a "conservatory" course of music study emphasizes that the learner follow verbal directions, decode written notation, develop his playing or singing techniques, and that he meticulously follow the directions of a conductor or a teacher. General intelligence is critical to such behaviors. Thus, it becomes questionable to teach to general intelligence to the exclusion of audiation. This seems all the more bizarre when it happens at the primary and elementary levels of learning. There is also a danger that the focus on decoding notation or being micro-managed by a teacher can serve as a substitute for one's own audiation, and perhaps even mask one's own audiation. After all, the students in this study were unaccustomed to responding in the manner required by AMMA.

## CHAPTER 5

## SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

## Summary

Purpose and Problems

The purpose of the study was to obtain and examine data from the Advanced Measures of Music Audiation (AMMA) to determine if AMMA is both a reliable test and a valid test of stabilized music aptitudes for college music majors. In addition, through the use of grade point averages in applied music, aural skills, and written harmony the predictive validity of AMMA was investigated.

The problems of the study were; 1) to establish the reliability of AMMA, 2) to establish the level of immunity of AMMA to practice, instruction, and maturation; and 3) to determine if AMMA scores could serve as a predictor of success in music study.

Design and Analysis

AMMA was administered on two occasions to the same 40 undergraduate music majors. The first administration provided a set of pretest scores, and the second

administration 8 months later provided a set of posttest scores. During the interim the students were engaged in applied music study, aural skills training, and written harmony classes.

Reliability of the pretest and the posttest scores was established first. Next, the differences between the pretest and posttest results were examined for any statistically significant differences. Students' mean grade point averages (GPAs) in applied music, aural skills, and written harmony were obtained. The results were correlated with the AMMA pretest Total raw scores to determine the strength of the correlation. How strong those correlations were would suggest the predictive validity of AMMA.

### Results

AMMA Total test raw scores reliability coefficients were obtained by the split-halves method. They were .82 for the pretest and .84 for the posttest. Both coefficients were close to the reliability coefficients reported in related studies. In addition, the means and standard deviations obtained in this study were close to means and standard deviations reported in related studies.

The differences between the AMMA pretest and posttest composite mean raw scores were examined for any statistical significance at the .05 level of confidence. The obtained mean raw score difference of -1.8 between the two test

administrations was not large enough to be statistically significant.

Correlation coefficients between the AMMA Total mean raw scores and the mean GPA for applied music, aural skills, and written harmony, were calculated. The coefficients obtained turned out to be slight to negligible, but when the top 20% and the bottom 20% of AMMA Total raw scores were examined, the results revealed stronger relationships.

### Conclusions

Advanced Measures of Music Audiation is a reliable test with split-halves correlation coefficients of .82 and .84. In addition, for undergraduate college music majors, it seems immune to the effects of music instruction. The test may be used to predict potential in music achievement for high scoring and low scoring students.

### Recommendations

AMMA needs to be used far more widely than it is currently used. Typically, all the college music majors in this study were screened for their academic achievement with established standardized tests but they were not screened for their music aptitudes. It seems ironic that persons who major in music do so without knowing their musical potential.

In addition, the music faculty were equally ignorant of their students' music aptitudes scores even though they would measure and evaluate their students' achievement in music.

An additional argument for the wider use of AMMA is the hope that through a heightened understanding of music audiation, as outlined on page 7, audiation may play a greater role in music instruction. For example, instrumental students could profit from singing and moving to music. It could result in music audiation approaching or even surpassing the role of general intelligence in how music is taught. That, in, turn could help music educators to reassess how they measure and evaluate their students' music achievement.

Underlying some of the disparities between testing for music aptitudes and how music is taught is audiation. Audiation is known and understood by only a minority of music educators. One can but strongly recommend to music educators the need for the profession to thoroughly investigate and discuss audiation.

## REFERENCE LIST

- Arms, K., & Camp, P. S. (1979). Biology. New York: Holt.
- Bartlett, J. (1980). Familiar quotations. (15th ed.) Boston: Little Brown.
- Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990, October 12). Sources of human psychological differences: The Minnesota study of twins reared apart. Science, pp. 223-228.
- Edelman, G. M. (1987). Neural darwinism. New York: Basic Books.
- Epstein, H. T. (1978). Growth spurts during brain development: Implications for educational policy and practice. In J. S. Chall & A. F. Minsky (Eds.). Education and the brain. Chicago: The Seventy-seventh Yearbook of the National Society for the Study of Education, Part 2.
- Estrella, S. G. (1992). Effects of training and practice on Advanced Measures of Music Audiation scores. Unpublished doctoral dissertation, Temple University, Philadelphia.
- Fullen, D. L. (1993). An investigation of the validity of the Advanced Measures of Music Audiation with junior and senior high school students. Unpublished doctoral dissertation, Temple University, Philadelphia.
- Gordon, E. E. (1965). Music aptitude profile. Boston: Houghton.
- Gordon, E. E. (1976). Learning sequence and patterns in music. Buffalo: Tometic Associates.
- Gordon, E. E. (1987). The nature, description, measurement, and evaluation of music aptitudes. Chicago: G.I.A..

- Gordon, E. E. (1988). Learning sequences in music (3rd ed.). Chicago: G.I.A..
- Gordon, E. E. (1989). Advanced measures of music audiation. Chicago: G.I.A..
- Gordon, E. E. (1990). Predictive validity study of AMMA. Chicago: G.I.A..
- Petitto, L. A., & Marentette, P. F. (1991, March 22). Babbling in the manual mode: Evidence for the ontogeny of language. Science, pp. 1493-1496.
- Seashore, C. E. (1919). Measures of music talent. New York: Columbia Phonograph.
- Seashore, C. E. (1967). Psychology of music. New York: Dover. (Original work published 1938)
- Seashore, C. E., Lewis, D., & Saitveit, J. (1957). Seashore measures of music talent. New York: Psychological Corporation.
- Shuter-Dyson, R., & Gabriel, C. (1981). The psychology of musical ability (2nd ed.). London: Methuen.
- Staddon, J. E. R. (1983). Adaptive behavior and learning. Cambridge, England: Cambridge University.
- Taylor, S. E. (1964). Listening, What Research Says Series, no. 29. Washington, D. C.: Association of Classroom Teachers of the NEA, 4.
- Wing, H. D. (1961). Standardized tests of musical intelligence (5th ed.). Sheffield, England: Sheffield City Training College.