

2021

### Edwin Gordon's Contribution: An Appraisal

Richard Colwell  
*Boston University*

Frank Abrahams  
*New England Conservatory*

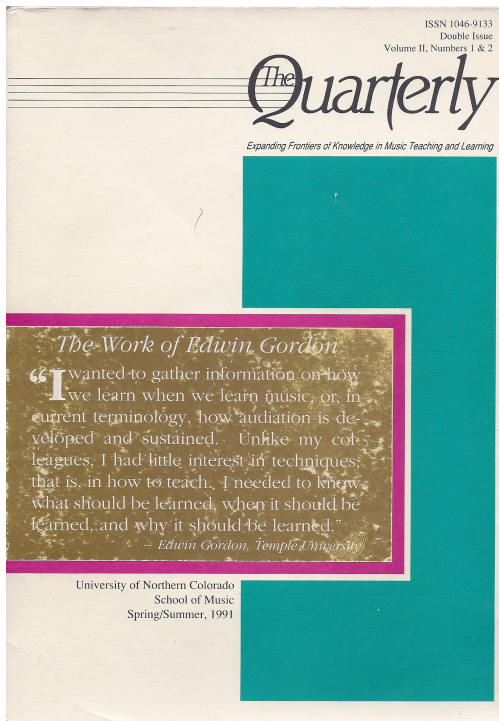
Follow this and additional works at: <https://opencommons.uconn.edu/vrme>

---

#### Recommended Citation

Colwell, Richard and Abrahams, Frank (2021) "Edwin Gordon's Contribution: An Appraisal," *Visions of Research in Music Education*: Vol. 16 , Article 6.

Available at: <https://opencommons.uconn.edu/vrme/vol16/iss2/6>



**Title:** Edwin Gordon's Contribution: An Appraisal

**Author(s):** Richard Colwell and Frank Abrahams

**Source:** Colwell, R., & Abrahams, F. (1991, Spring & Summer). Edwin Gordon's contribution: An appraisal. *The Quarterly*, 2(1-2), pp. 18-36. (Reprinted with permission in *Visions of Research in Music Education*, 16(2), Autumn, 2010). Retrieved from <http://www-usr.rider.edu/~vrme>

*Visions of Research in Music Education* is a fully refereed critical journal appearing exclusively on the Internet. Its publication is offered as a public service to the profession by the New Jersey Music Educators Association, the state affiliate of MENC: The National Association for Music Education. The publication of VRME is made possible through the facilities of Westminster Choir College of Rider University Princeton, New Jersey. Frank Abrahams is the senior editor. Jason D. Vodicka is editor of the *Quarterly* historical reprint series. Chad Keilman is the production coordinator. The *Quarterly Journal of Music Teaching and Learning* is reprinted with permission of Richard Colwell, who was senior consulting editor of the original series.

*“The stimulation of a profession comes from wrestling with ideas, from the process of sorting out the claims made by the best minds in the profession.”*

*--Richard Colwell and  
Frank Abrahams*

**Section 2:**  
*Three Perspectives on the Work of  
Edwin Gordon*

# Edwin Gordon's Contribution: An Appraisal

---

**Richard Colwell**

*Boston University*

**Frank Abrahams**

*New England Conservatory*

**E**dwin Gordon's ideas stand at the opposite end of the spectrum from those of music education's philosophers and from the concerns of many researchers. Thus, it is important to critically examine Gordon's ideas as revealed through his tests, methods of instruction, and texts. Also, in this review, we contrast Gordon's ideas with those of Bennett Reimer, another leading figure in music education.

Unfortunately, there is no research which is sufficiently definitive to indicate the degree of truthfulness or error in the research of Gordon and his writings. There is likely some truth in all approaches to the teaching and learning of music, and our objective is to emphasize the need for the profession to wrestle with ideas and to more fully explore research that assists in discussion making.\*

Gordon is a behaviorist. He believes that learning progresses through successive stages, and he uses psychological principles to explain music learning. He has not accepted the cognitivist philosophy of learning, and he stresses the importance of each student achieving his or her full potential. Gordon is less concerned with the social goals of music which would raise the general education level of music education to minimum standards. Also, there are few references in his writing about group musical goals, although some goals are accomplished through group processes. In this article his

ideas will be explicated along with a discussion of some of his major ventures in music education.

In spite of the promises and problems of Gordon's work, he remains the consummate musician. His performing abilities are impeccable, and he has had considerable

---

*“Gordon continues to this day to persist in the belief that knowing a student's aptitude is a critical element in structuring instruction.”*

---

professional experience in the jazz field. As with each of us, his ideas may stem from his own experiences with music and his recollections about how he learned. No doubt he was a fast and thorough learner, and his expectations may reflect his own competencies in music.

## Gordon's Tests

Gordon has devoted his professional career to investigating the role of music aptitude as it affects music instruction. Beginning with his own doctoral dissertation in 1958, Gordon has focused his research efforts on identifying those elements and processes that impact upon a student's progress in learning to “hear” music. Gordon has carefully developed a music achievement test and a series of music aptitude tests, investigated the limits of aptitude measurement, and most recently has developed instructional sequences in instrumental and vocal music that are based upon nearly two decades of research. In some respects, his work is

---

\*In addition to the review of Emily Brink (1983) which appeared in No. 75 of the *Bulletin of the Council for Research in Music Education*, the material that the senior author prepared for test reviews in *Mental Measurement Yearbooks* (1972) and *Test Critiques* (1986) has been used in this article. Research on some of Gordon's tests was conducted specifically in preparation for this article.

rather specialized, and a student taught by the “Gordon system” is more likely to demonstrate achievement on one of his tests.

The flaws in his work are often a result of Gordon’s attempt to “teach” the reader or user of his research rather than carefully and objectively presenting and interpreting the data. Even in the manual for his stabilized music aptitude test, the *Musical Aptitude Profile* (MAP) (1965), Gordon devotes pages to suggestions about what should be done by the teacher for the student who scores below average on the test. If there is one student in a school whose score is less than it might be because his or her aptitude has not stabilized, Gordon wants the teacher to begin an instructional campaign to change that student’s score before it’s too late.

### ***Musical Aptitude Profile***

The publication of Gordon’s *Musical Aptitude Profile* (1965) more than a quarter of a century ago forced the profession to admit the importance of this music educator and his ideas about talent and instruction. The test was well crafted and, most importantly, it was accepted for the catalog of Houghton-Mifflin, a prestigious publishing firm with an excellent reputation in testing. A number of important research projects were occasioned by the appearance of the *Musical Aptitude Profile*.

To establish the validity of the test, Gordon decided that all of the students in classes at a selected grade level should be involved; his question involving the music aptitude of the population, not the music aptitude of those students who had elected to study music. This insightful decision marked a stunning advance in the profession at a time when all previous research on music aptitude had been conducted with selected students. If the question to be answered is, “Among the students who have indicated an interest in instrumental music, what are their levels of music aptitude?” then the traditional technique of establishing validity remains the more appropriate method, as the validation research has also been conducted with students demonstrating an interest in music. If the question is, however, “Among the sea of bright and shining faces, which students are the more talented and how can that

information be used in selecting students for enrichment opportunities or in individualizing instruction within classes of general music?” then the teacher/researcher would want data representative of all students.

Experienced teachers know there are few absolutely right or wrong answers in education, thus requiring us to know how to interpret data for our own situation. Gordon’s validation of his aptitude test was an effort aimed at establishing the aptitude of all students attending the schools. His aptitude test, the MAP, was carefully constructed over an eight-year period.

Gordon carefully and thoroughly investigated most aspects necessary for a valid research design. He relied heavily on correlation techniques, and a correlation study is vulnerable to several criticisms. One correlation has been of special interest to Gordon, and that is the relationship of his aptitude test scores with measures of intelligence.

Gordon does assume that instruction aids achievement, and for establishing validity he relies on the relationship between a student’s aptitude and his or her achievement in instrumental music. His rationale is that the student with the higher aptitude will obtain a higher score on an achievement test than a student with lower aptitude.

Gordon’s selection of instrumental music is understandable. For one reason, the objectives of fifth grade music vary from community to community; standards of fifth grade (general) music vary, the amount of instruction varies, student motivation, interest, and commitment varies, and there are few valid measures of achievement for fifth grade music classes. Thus, Gordon’s use of instrumental music achievement as the criterion was clever.

To obtain the cooperation of every child, Gordon provided the students with new instruments. Gordon may have been influenced by Professor Schneider of the University of Northern Iowa, as Schneider had developed a successful program where every third grade student was required to study violin for one year. Gordon, however, chose to use band instruments and approached the National Association of Band Instrument

Manufacturers to make a substantial gift of instruments to the University of Iowa. One assumes that the band instrument manufacturers were interested in the study because a valid aptitude test would strengthen recruiting efforts for instrumental music programs, and, if the recruiting were done more intelligently, there might be fewer dropouts, the bane of public relations efforts in instrumental music.

A review of the objectives for the aptitude test reveals that four of the five objectives are sympathetic to the cause of instrument manufacturers:

1. to identify musically talented students who can profit most from and contribute most to school music activities;
2. to aid in the formulation of educational plans in music;
3. to compare the collective musical aptitudes of groups of students; and
4. to apprise parents of the musical aptitudes of their children (Gordon, 1967).

The remaining objective is "to adapt music methods and materials to the individual needs and abilities of students by compensating for their specific music weaknesses and by enhancing their specific music strengths" (1967, p. 1).

The validity study was conducted with five midwestern school systems. The classrooms were apparently not randomly selected within the cities. During the study, Gordon allowed the students to study privately and to participate in elective school and community musical activities. One can guess that private lessons and the extracurricular activities were not spread over the continuum but were elected by the better students. We expect this fact of additional instruction to influence the achievement test scores, especially those related to performance. Gordon does not provide data separately for those students studying privately, but the correlation of aptitude and private instruction is .04, private lessons and achievement .15. The relationship of achievement to 19 of the 23 factors ranging from sex to mother-attended-college was greater.

One must ask if Gordon's achievement test was a valid indicator. Correlations between instrumental teachers' ratings of achievement and the *Musical Aptitude Profile* were also

low. With the composite test score, they are .37, .39, and .35 for the first, second, and third years respectively. The reader should wonder about this low correlation between the teachers' rating of achievement and the predictive ability of the *Musical Aptitude Profile*. In order to document the power of the MAP, Gordon selected the upper and lower ten percent of the students (as measured by his test) and compared the mean scores of these two groups on the various achievement tests and teacher ratings. He obtained a difference between these extreme groups but he also should have conducted a similar study using the top and bottom ten percent of the students as identified by teacher ratings.

Gordon's explanations for the low correlation with the teachers' ratings include that the teachers' rating may be weak in validity, that extramusical factors may influence teachers, or that teachers have a lack of ability to discriminate among students with similar but not equal talent (Gordon, 1967, p. 18). As Gordon is a pioneer in careful music research, one is not prone to disregard his conclusions. Most music educators, however, think that a competent teacher working with the same children constantly on instrumental instruction over a period of three years would be better able to judge musicality and music achievement in a child than an impartial judge making a judgment based on two hearings of three etudes. The possibility exists that Gordon's objectives for the aptitude test and for the achievement test were similar.

Gordon is apparently uneasy with using teacher ratings as valid measures of music achievement. The correlations between composite test scores and teachers' estimates ranged from .64 to .97, an impressively high correlation. With a correlation of .97, an aptitude test was not needed; the teacher's estimate is as accurate a predictor as that provided by the results of the profile. A correlation of .64 is also not too shabby; these correlations are sufficiently high to provide assurance to instrumental teachers of their professionalism. When these teachers used a student's performing ability on etudes composed by Gordon as a criterion, the

correlations dropped to a range of .12 to .55, an indication that teachers have more difficulty evaluating achievement than aptitude. These teachers were affected, as we assume they should be, by students' effort, progress, and personal factors as they judged the students' performances. Gordon, on the other hand, suggested that the relatively low correlation might be due to the high selectivity of students enrolled in musical performance or the unreliability of the performance measures. Admittedly performance groups are selective and somewhat homogeneous, but few groups are sufficiently selective to reduce the range of achievement to such an extent as to substantially affect reliability.

Gordon (1967) states "All aptitude tests are, through necessity, to some degree achievement tests, just as all achievement test scores to some degree reflect aptitude, depending, of course, upon the extent to which the effect of achievement on performance has been successfully minimized or maximized in relation to the effect of aptitude" (1967, p. 5). This statement is true, but misleading in that the relationship is not of the same magnitude. Achievement tests are substantially affected by one's aptitude and are designed to measure one's aptitude plus the impact of instruction, practice, motivation, and interest (or sometimes the determination of the parent to have success). The developer of an aptitude test, on the other hand, must go to considerable lengths to minimize the influence of instruction, environment, interest, and other factors that have been shown to be related to achievement.

Gordon's three-year validity study was conducted in fourth and fifth grade classrooms in Davenport, Cedar Falls, Iowa City, and Racine. These towns are not typical communities, and the interest of parents in the music program would be greater than in many other communities. Although selectivity causes many known and unknown problems, one might expect students from these four communities to have "tried harder" to have been more accepting of the three-year commitment in trade for a shiny new instrument. On the other hand, requiring a student to study instrumental music may be no different from requiring a student to study general music except for the expectations for

instrumental music and the consequences of mixing motivated and nonmotivated students in the same instrumental class lesson.

The criteria for success were adjudicator ratings of melodic, rhythmic, and expressive aspects of tape recorded performances of short etudes prepared in advance with teacher help; the same etudes prepared in advance without teacher help; and the etudes sight read. In addition, teacher ratings of each student's musical progress were compared to other students in the group and their performance on an objective musical achievement test. The N was 241 students in eight classrooms in five different schools in the four cities.

By the end of the third year, the reliability figures had increased and prediction strengthened. (Reliability would be greater because the spread within the group would have increased.) The students who had studied privately for three years would differ from the students who had developed a dislike for instrumental music and were less than cooperative during instruction and testing. Split half reliability is greater with heterogeneous than with homogeneous groups. Reliability is also enhanced by the length of a test; the *Musical Aptitude Profile* is a lengthy test requiring nearly three hours to administer. The predictive validity increased from year 1 to year 3, indicating that what is learned in music that is related to aptitude takes time to develop. Experimental students received private and small-group instruction on instruments. We conclude that the predictive validity of success from this test for a general music class that meets only once a week would likely be lower.

Gordon's interest in instruction had appeared by 1967 (Gordon, 1967, p. 34). "In this connection, it is interesting to speculate on the possible effect upon *Musical Aptitude Profile* predictive validity if the teachers could have known their students' initial aptitude scores—thus making it possible for them to adapt instructional procedures more effectively from the very outset of the study" (1967, p. 34). Gordon continues to this day to persist in the belief that knowing a student's aptitude is a critical element in structuring instruction.

Gordon is correct in his manual that the

achievement scores were affected (lowered) as a result of the instructional situation. Being assigned an instrument which the student grew to dislike along with the normal physical changes of growing up are important factors but these factors should be constant across aptitude levels and have minimal effect on the correlations. Gordon does not provide a hypothesis for the lower relationship between aptitude and achievement when the student received instruction than occurred without instruction. Could instruction have made a difference, causing the groups to become more homogeneous in their achievement, and resulting in lower correlations?

The scores of the students who lost all interest in music ( $N = 23$ ) were below the mean on the MAP but not as low as one might expect. Total group  $N$  was 46. The students who lost interest obtained a score of 42.5, a low mean score but far above the 33.7 mean score scored by the lowest 10 percent of the students in the study, many of whom did not lose interest.

Gordon (1967) argues for use of the composite MAP score because it is "clearly significantly better for predicting success in music than any single score provided by the test" (p. 39). This admonition must be taken into consideration when using IMMA and PMMA, tests that are shorter and lacking the sensitivity section which is part of the composite score.

Instrumental instruction did not affect scores on the MAP. "It appears that formal instruction in instrumental music does not have any appreciable effect upon *Musical Aptitude Profile* scores" (1967, p. 40). The finding of lack of impact of instruction is contradicted in Gordon's later writing about developmental aptitude and stabilized aptitude. Although aptitude test scores are not significantly correlated with amount or practice, they are related to a student's attitude toward practice (1967, p. 44) and related to attendance at concerts, summer lessons, and extra school activities. Additional research on this topic is needed.

In his three-year validity study, the relationship of teacher's rating of success with Gordon's achievement test was low. Teacher composite ratings of progress correlated .34

with musical achievement. The correlations were .26, .48, .28, and .47 in four schools. The validity of Gordon's achievement test is critical to the arguments in this article, as his achievement test was used as the basis for constructing the *Iowa Tests of Musical Literacy* (1970). The tonal and rhythm patterns in the ITML are the basis for Gordon's thinking about sequencing and the exercises selected for *Jump Right In: The Music Curriculum* (Gordon & Woods, 1986). Musical sensitivity correlated .61 with composite achievement, a correlation as high as any other on the battery. Rhythm imagery correlated .61 and tonal imagery .57 to musical sensitivity, indicating that sensitivity contributed measurably to the composite score of music aptitude.

At the time that he was developing MAP, Gordon claimed that his primary interest was improving instruction. Gordon believes that he could not compare the effectiveness of two instructional methods without knowing the student's aptitude because a researcher might mistakenly reject a method; the apparent failure of the method might be due to lack of talent in the subjects.

In MAP, the mean item difficulty for questions on rhythm with students in grades ten to twelve exceeded normal extreme limits of .80 for a norm-referenced test (.84). Even the total mean item difficulty of .79 for R1 and .73 for R2 is surprisingly high. The reliability indicates that students can successfully answer most questions, as all of the sensitivity items have a difficulty index of .70 and above.

The discrimination indices are higher than would be expected with such easy questions. Either there is an error in the manual or the test could be shortened considerably without loss of validity.

Gordon has not continued his interest in stabilized aptitude since the data on MAP were gathered. His interest is now focused on developmental aptitude, a term that can be interpreted as early achievement.

### **Primary Measures of Music Audiation**

Gordon's *Primary Measures of Music Audiation* (PMMA) (1979) comprises a music aptitude test for kindergarten and primary-grade children. The test consists of 80 pairs

of musical stimuli, 40 tonal and 40 rhythm patterns. Electronic sounds generated by a synthesizer, somewhat similar to the sounds generated for the older *Measures of Musical Talents* (Seashore, 1960), are used as stimuli. The task for the student is to determine whether the paired patterns are the same or different. Tonal patterns consist of musical phrases two to five tones in length with 13 two-tone patterns, 25 three-tone patterns, and a single four- and five-tone pattern. Patterns are voiced in the middle register and are based on tones from tonic, dominant, and subdominant chords. The tonal patterns are in major or minor tonality (no other modes are used). The two phrases are in the same key tempo, and the metrical and rhythmic elements are identical. Although Wehner (1985) indicates that a discrepancy on three items with the keyed answer, the test is correctly keyed.

Rhythm patterns range from two to seven tones, all sounded on the same pitch. Most patterns are four to six notes in length. The rhythm patterns are almost devoid of silence—only one item contains a rest, and it is a 30-second rest; thus students will likely not perceive this rest as having a specific duration in the rhythmic patterns test.

The patterns used in the test are drawn from those Gordon found to be the easiest in an earlier study in which his concern was the development of a taxonomy of tonal and rhythmic patterns. Whether the selection of easy or difficult patterns makes any difference in measuring aptitude is unknown; the length of the pattern is presumed important.

The PMMA is designed for kindergarten, first, second, and third grade students, and all four levels of students were involved in the test development. One study used fourth grade students, but these subjects were found to be too old to provide valid and reliable data. The test items are not overly difficult; kindergartners average 64 percent correct answers on the tonal test, a figure that rises to nearly 86 percent for third-grade students. The rhythm test is slightly more difficult, with average difficulty indices ranging from .58 for kindergartners to .75 for third-grade students.

The PMMA can be administered and scored by the music teacher. The instructions are

clear and professional with minute attention to detail. The author insists that the tonal pattern test must be given first, although no empirical reasons are given for this ordering. Gordon's explanation is that students will better understand the testing task if they begin with the tonal patterns. Because computer scoring is not feasible for pictographic answer sheets, scoring of the test is done by hand. Templates allow the teacher to see if students have circled all of the answers or have patterns marked. Although a composite score is computed, Gordon's primary interest is that the examiner interpret test results using two scores, one for tonal patterns and one for rhythm patterns.

Four school systems provided the data in the development of the *Primary Measures of Music Audiation* (1979); the norms presented in the manual are from one school system in upstate New York (N = 873). One might more reasonably expect this sampling scheme in the development of a criterion-referenced test than in the development of an aptitude test.

Gordon discusses content validity, concurrent validity, and congruent validity; however, he provides little data for any of these. For content validity, Gordon suggests that the PMMA measures keyality and tempo within the tonal and rhythmic pattern discrimination. He accepts at face value that tone and rhythm are the two major components of aptitude; however, in his tonal and rhythm pattern research, the "ability of fourth grade students to hear pairs of patterns as being the same or different was found to have virtually no correlation with stabilized music aptitude as measured by the *Musical Aptitude Profile*" (1981, p. 73). This statement could be interpreted to mean that there is almost no relationship between developmental aptitude and Gordon's definition of true music aptitude. The same data (1979, p. 90) are used, however, to argue that scores from the PMMA are related to the MAP:

The correlations between the *Music Aptitude Profile* and the *Primary Measures of Music Audiation* are comparable to the longitudinal predictive validity coefficients reported for the *Musical Aptitude Profile*. Although not conclusive, thus far these data offer the strongest objective evidence of the validity of the *Primary*

*Measures of Music Audiation* (p. 90).

The data supporting the argument that Gordon is not measuring aptitude with the PMMA are more convincing than his concluding sentence in the manual stating that he is.

Concurrent validity is dismissed because of the difficulty of distinguishing between informal and formal music achievement on the one hand and innate music aptitude on the other (1978, p. 86). Our point is that there can be no reliable criterion. If Gordon's statement was true, criterion-related validity would not be possible for any aptitude or intelligence test. The construction of all aptitude tests poses the same problem. Gordon focuses on the small correlation between PMMA scores and test scores on the *Lorge-Thorndike Intelligence Tests* (1954-1956) *Stanford Achievement Test 1982*, and *Metropolitan Readiness Test* (1987). Although of interest, this negative argument is not convincing as an indicator of validity. The case for congruent validity rests on the congruence between the test scores of 227 fourth-grade students on his two tests, PMMA and MAP.

Gordon is to be commended in providing both internal and test-retest reliability data. Internal reliability is especially high. His logic regarding the test-retest data is that a correlation of .60 between two administrations of the test a week apart seems good (high) to him, and a correlation of .51 between two parts of the test seems low—he argues that only 25 percent of the variance is common. The correlation between the two parts of the test of about .50 was confirmed by our administration of the test, and similar figures were found in at least two other studies. Additional care is needed in drawing conclusions from these data. The test-retest reliability computed by us using students in first and second grades was .61. The interval was one year. The combined grades contributed to a speciously high reliability; a test-retest reliability of .5 for a single grade is probable.

Of most interest to the technician is Gordon's item analysis table for the two subtests (see Table 5, 1982). The items are shown to be very easy even for first-grade

children. Twenty-four of the 40 items in the tonal test had a difficulty index of .80 and above; the average difficulty index for third-grade is 85.6, an indication that most students obtain the correct answer. Computing his item discrimination by point biserial (the presumed method, as Gordon does not so state) provides impressive discrimination values. These unlikely, but possible, figures occur only when a test resembles a Guttman rating scale. Gordon's statement that "there is a considerable range of item difficulty levels for each test in all grades" (1979, p. 70) cannot be taken seriously; one difficult question (#19) hardly gives "considerable" range; it is the only one with a difficulty index of less than .5 for third grade students. The point biserial formula used is not so much a measure of discrimination between those who know and those who do not know as it is a measure of the relationship between the individual item and the total score. These data, as suggested earlier, tend to explain the results Gordon obtained in his factor analysis as well as some of his other statistical data.

Gordon's primary argument for the validity of the PMMA, as it was in the MAP, is the low correlations with other school tests. For example, he indicates that scores on the PMMA correlate around .3 with Stanford's test in mathematics and reading. The basis of the argument is subtests. The result of factor analysis is that the primary factor identified is whether the student answers "Same" or "Different." Gordon's item-analysis data indicate the probable futility of obtaining insights from this factor analysis to which considerable manual space is devoted.

Gordon seems reluctant to drop or change an item or item type once he has formulated his testing hypotheses, which fortunately are usually based on extensive and careful thought. An example of this penchant for stasis can be found in the PMMA, where all the tonal items were retained throughout the test's development. A few rhythm items were changed because of that subtest's lower and less reliable score, but Gordon claims that even this revision had little effect on validity or reliability—an argument to his audience that he might just as well have stayed with the original rhythm items (Gor-

don, 1979, p. 19). It is important to note that Gordon operates within a single philosophy of music instruction and learning and that he makes every effort to be consistent.

The author repeatedly argues that the test is valid only during that period when the young child's music aptitude is in a state of flux. Once musical ability stabilizes, which Gordon believes is around the age of 9, the data from this test are no longer valid; the data are not valid even for primary age children if their aptitude has already stabilized. A major use of the test results for Gordon is that scores from the tests should serve a diagnostic function based on whether a student is in the high, middle, or low group on pitch and rhythm. This suggestion, however, has limited value, at least with third grade students and perhaps with others also.

The test functions as a criterion-related achievement test with a ceiling that is rather easily attained. Missing a question lowers the percentile rank of third grade students by six percent. Thus a student who misses five questions on the rhythm test is at the eighty-first percentile, six questions at the seventy-fifth. Teachers with strong music programs will find the norm tables for third-grade students to be disconcerting and may believe that their students have not scored well, when in fact the ceiling effect is influencing their scores and their class rank.

In administering the test in two different communities, we obtained scores even higher than the already "high" scores obtained by Gordon for his norms. Fewer than 15 percent of our sample scored below 35 on the tonal test. The table of norms in the manual suggests that the test is inappropriate for third grade students (too easy), but these are likely underestimated; the norms are most appropriate for first and second grade students. Accordingly, the test is not appropriate as an aptitude test for all the recommended ages.

No review of Gordon's work is complete without documenting his primary concern for instruction, not for testing. Pages 50-64 of the PMMA manual (Gordon, 1979) are devoted to suggestions for teaching. Gordon is a good teacher, but he does make unsubstantiated statements in what is presented as a scholarly test manual:

"Unconscious listening to music is most important before the age of three" (p. 51).  
 "Instrumental music should be played most of the time" (p. 52).  
 "Parent or teacher singing to a child is of equal importance (to instrumental music)" (p. 52).  
 "Rote singing probably contributes less than spontaneous singing to the development of a child's tonal aptitude" (p. 53).  
 "There is direct correspondence between the quality, quantity, and diversity of rhythm activities of a preschool child and rhythmic aptitude" (p. 54).  
 "Large scale muscle movement should be encouraged" (p. 55).  
 "There are two types of nonsingers; two types of out of tune singers" (p. 57).

He describes how to teach a rote song and much more. These (pedagogical) statements are often logical, but none is supported by research.

### ***Instrument Timbre Preference Test***

Gordon's *Instrument Timbre Preference Test* (1986) is designed to provide information to teachers, students, and parents that will be helpful in selecting a musical instrument for the student. The data from the test are to be used in aiding the teacher in the prediction of successful performance on a wind instrument by beginning music students.

The task for the examinee is to indicate the preferred timbre when two different timbres are sounded by a Moog synthesizer. Seven different timbres are used in the test. Each of the seven is paired twice with every other timbre, with each being heard once as the first item of a pair and once as the second item. Forty-two items constitute the test. The duration of the tape is 22 minutes; thus, administering the test takes about 30 minutes. The timbres generated by the synthesizer match the timbres of musical instruments only in a general way. Gordon's rationale was to construct a test where the cultural association a student has with a particular instrument or performer is minimized. The timbres are intended to be close to those of the wind instruments of the band and orchestra.

The philosophy behind this test is that a student will be more successful on an instrument that emits a timbre pleasing to him or her. Gordon suggests that use of the

data will reduce the unknowns in instrumental music education. He believes that scores on the MAP account for about 56 percent of the reason for success in the study of instrumental music, and that intelligence accounts for as little as four percent; further, the timbre test accounts for an additional ten percent of the reasons for success in instrumental music. Gordon adds, "It has been found that the timbre of the music instrument a student plays is second only to his music aptitude as an important factor in instrumental music" (Gordon, 1986, p. 5). We know of no confirming research for this statement.

The data to answer questions satisfactorily about the practical use of this test, unfortunately, do not exist. For example, we have no data on the education issues involved with encouraging nonvolunteers to enroll in instrumental music whether on appropriate or inappropriate instruments. Of practical importance to teachers is the ability to provide feedback to all students following testing. After scoring the test seven times, data for only the extreme high-scoring students and extreme low-scoring students are the only useful data.

Gordon's data, gathered at the time the test was developed, indicate that students who were aided in the selection of their instrument by the results of this test actually attained a higher level of skill in the performance of technical exercises. The causal relationship is unclear, because these same students were less interested in participating in school ensembles. The test, of course, is designed as a supplement to a music aptitude test and has limited value as a stand-alone item. If the test results contribute ten percent to the variance in performance scores, Gordon's inference is that this ten percent is related to a "factor" of preference for timbre and is not simply an additional part of general musical aptitude. If his inference is true, the test would be valuable in its own right and should provide information, given the student's talent, that he or she would be better advised to select an instrument with the preferred timbre.

The test is to be given at whatever age instrumental music instruction is begun. The author suggests this is to be fourth, fifth, or sixth grade. Scores of 10, 11, and 12 for a

particular timbre are favorable scores, and the student scoring ten and above should be encouraged to select an instrument with similar timbre. Scores of three and above or nine or less are ignored, as the student has no strong preference one way or the other. Scores of 0, 1, and 2 toward a timbre indicate a dislike for that sound, and thus the instrument should not be selected.

The technical data in the manual that are of value to the user are found in the reported test-retest reliability. After one week, reliability is reported to be about .70. Of importance, however, is the reliability of individual timbres. When the rationale of the test is based on the concept of reliability, the student must respond consistently throughout the seven parts of the test in order to score high enough or low enough for a preferred timbre to be determined.

The data from one school cited in the manual provide mean scores for each of the seven timbres for students in grades three through eight. Preference in this sample appears to be age related, not reassuring information for the parent or teacher who hopes that a student will select an instrument and stick with it long enough to develop some useful skills. In his sample, more students prefer woodwind than brass.

Gordon is a careful worker in all of his research; the development of the *Instrument Timbre Preference Test* took several years. To establish criterion-related validity, which is the relationship with success in instrumental music, Gordon developed and administered the test, allowed the students at least one school year to learn an instrument, and assessed their achievement in relationship to their initial timbre test scores. Results showed that students playing appropriate instruments (i.e., the preferred timbre) received consistently higher ratings on the tone quality dimension of the assessment.

Criterion-related validity is provided by combining the scores attained by performing three etudes and comparing these combined scores of an experimental group ( $N = 33$ ) and a control group ( $n = 47$ ). Performance ratings on the etudes were based on the tonal, rhythm, tone quality, and expressive aspects of the students' performance with each of the three exercises sight read. A

score of 120 was possible for all three exercises; the mean scores were 46.4 and 37.4, both with rather large standard deviations (20.1 and 17.1 respectively). With no right or wrong answers, the test is neither a criterion- nor a norm-referenced test.

We administered the *Instrument Timbre Preference Test* to a sample of the same size as that used by Gordon in obtaining his initial data. Of the 77 students who began instrumental music study (in one of two school systems used), only eight selected the instrument recommended by the test's results. Four students selected instruments that, by their timbre, should have been avoided. Of the eight "matches," five were for flute timbre. Possibly these students were starting on the wrong instrument, but the teachers interviewed indicated that they would rather have the student begin on a student-selected instrument (presumably an instrument the student will play in band or orchestra) than to have the benefit of a slightly higher score on a performance evaluation. Admittedly, many of the students used by these reviewers may drop out of the school music program, but Gordon also does not have definitive long-term dropout data on his sample. Fifty-two percent of Gordon's experimental group discontinued study after one year. Precise data are not available, but a 52-percent dropout rate in the first year would seem roughly equal to the norm.

Several nagging questions remain. Race, for example, seemed to be a factor in determining preference for the artificial sounds generated in the test. For instance, blacks liked the tuba sound, but seldom preferred the other brass instruments. In addition, of 87 college music majors tested at the University of Illinois, none selected the timbre of their major and 50 percent indicated a dislike for that timbre. In testing at the New England Conservatory, however, 18 of 22 students tested did select the timbre of their majors.

The primary contribution of Gordon in constructing the *Instrument Timbre Preference Test* is to raise substantive questions. If long-term preference and motivation to practice an instrument are not being measured, what is? To what extent is a synthe-

sized timbre transferable to the timbre of an instrument performed by a professional? Data exist that preferences are present and necessary when comparing a poor clarinet sound with a good one, or comparing good sounds in an instrument's extreme low register with sounds characteristic of extreme high-register pitches. These types of preferences appear to be important for success on an instrument. Presently, the results obtained from the use of the *Instrument Timbre Preference Test* will contribute more to knowledge within the research community than to teachers, students, or parents.

### Methods of Instruction

In several articles and books, Gordon has reported his thinking regarding teaching methods, by which he means "the sequential way in which immediate specific objectives are introduced in a course of study as they relate to the accomplishment of long-range comprehensive objectives" (Gordon, 1984, p. 8). For Gordon, the long-range objective for a music education program is music literacy; that is, the ability to read and write music with understanding. Understanding requires "notational audiation" (the ability to aurally imagine what one sees notated), which must be preceded by basic "audiation" (to be able to aurally imagine music through recall or creation). "In order to read and write music meaningfully, one must be able to hear music seen in notational form before it is performed, and to hear what one is composing" (Gordon, 1984, p. 3). The specific method Gordon promotes to achieve that goal relates to the sequential introduction of specific tonal and rhythmic patterns. To that end, he has worked on developing taxonomies of rhythmic and tonal patterns and theories to support them.

Rhythm is defined as comprising three elements: "tempo beats," "meter beats," and "melodic rhythm." Tempo beats (the basic pulse in a given piece) and melodic rhythm pose little difficulty. But meter beats introduce two essentially contradictory uses of the term "meter." Meter is commonly understood to refer to groupings of pulses, or tempo beats, into a larger temporal unit of fixed and steady duration. But Gordon

believes the opposite and defines meter (which he uses somewhat interchangeably with "meter beat") as consisting of division of the tempo beat, or pulse. Therefore, though 2/4 is considered duple because the pulse is divided into twos, 6/8 is considered to be in triple meter, not the traditional compound duple, because the pulse is divided into three meter beats. Gordon replaced the terms "tempo beat" and "meter beat" with "macro beat" and "micro beat," respectively, because of the confusion generated by the older terms (1971, p. 89). Their definitions, however, remain identical. Whenever Gordon refers to duple or triple meter, he refers not to meter signatures which account for a commonly understood duple or triple grouping of pulses, but to the division of pulses.

There is no argument among theorists that the traditional classification is inconsistent and unable to deal adequately with much music. Traditionally, the 3 in 3/4 meter refers to pulse, whereas the 6 in 6/8 usually refers to pulse division so that 6/8 could be notated as 2. But Gordon insists that a measure of 3/4 is actually felt as one half of a larger unit. "Most often the meter signature 6/8 indicates two beats of triple and not three beats of duple" (1971, p. 83). The key to understanding Gordon's definition of meter and his subsequent classification lies in the following statement: "Regardless of notational practice, young children usually subjectively organize tempo beats in pairs" (1971, p. 81).

Gordon does not support his statement with reference to studies, but more significantly bases his entire classification on the assumption that pairs of pulses are normative. Since there is essentially no theoretical room for a triple grouping of pulses, Gordon finds the traditional duple and triple classification of no use. All groupings are duple anyway. Therefore he appropriates the term "meter" for pulse divisions rather than pulse multiples. How then does Gordon account for patterns in 3/4 or 9/8? In 1971, they were assigned last place in his classification and named "uncommon unusual."

Gordon distinguishes two types of meter: "usual," that is, the tempo (macro) beats in pairs and of equal length, and "unusual." Unusual meter before 1980 was described as tempo beats in more than a pair, and/or of unequal temporal length, and/or without meter (1977, p. 53). At this time there was, therefore, still theoretical room for 3/4 and for 9/8. But in 1980 the definition of unusual meter was tightened to even more emphatically exclude 3/4 and 9/8 from his system. Unusual meter in 1980 was defined as groupings of macro beats which are of unequal length (1980, p. 92). Also in 1980, Gordon expanded the rejection of hearing in units of threes to include larger structural units. "In regards to form, a listener pairs not only beats but also measures and phrases (symmetrical and asymmetrical), sections, and even movements. In a sense, every whole becomes half of a new part" (1980, p. 88). The issue is one of levels of hearing. The logical extension would require all form to be essentially binary.

The term macro beat can no longer be limited to the concept of pulse; fundamental to the definition is the grouping of pulses. For example, in usual meter the pattern 2/4



contains two macro beats, but the pattern in Figure 1 contains only one, since 9/8 is audiated as half a measure of 6/8. Yet in unusual meter the pattern 7/8



is considered to contain three macro beats (1980, p. 93). What has happened is that Gordon moved from a definition of rhythm which began with the concept of pulse and ended up with a definition which is depend-

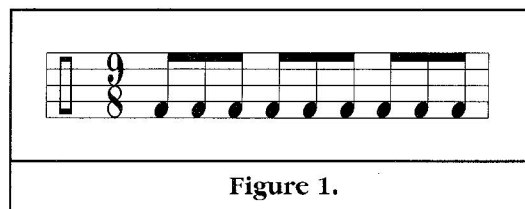


Figure 1.

ent upon meter. In 1980, his very discussion of the fundamental elements of rhythm is found under the headings “usual meter” and “unusual meter.” And all examples of macro and micro beats are given in metric context:

The most important distinction between usual meter and unusual meter is that micro beats are the meter beats in usual meter, and macro beats are the meter beats in unusual meter (1980, p. 92).

No where does Gordon provide musical examples to illustrate the application of rhythm patterns in the study of music. The brief examples in Figure 2 should suffice to demonstrate the problems of attempting to apply his terminology. The four notated measures of “Beautiful Dreamer” represent two pairs of macro beats, with one macro beat for each measure. Each measure in turn

contains three micro beats. But the basic beat is the macro, not the micro beat, since the beats must exist in pairs. “Down in the Valley” is notated in 3/4. According to Gordon’s system, each measure would similarly represent one macro beat, and two measures of 3/4 would actually be audited as one measure of 6/8. In this example, however, three measures of 3/4 are musically grouped together, forming recurring units of three macro beats, which are not permissible according to the theory of pairing macro beats. Similarly, the more complex opening measures of the third movement of Mozart’s “Fortieth Symphony” group three-bar units of 3/4. Cooper and Meyer (1960) refer to the difficulties of attempting a theoretical treatment of rhythmic grouping:

Rhythmic grouping is a mental fact, not a

Figure 2.

physical one. There are no hard and fast rules for calculating what in any particular instance the grouping is. Sensitive, well-trained musicians may differ. Indeed, it is this that makes performance of a piece of music possible.

Furthermore, grouping may at times be purposefully ambiguous and must be thus understood rather than forced into a clear, decisive pattern (p. 9).

In 1976, Gordon expanded his assertion that duple groupings of pulses are normative and included not only children, but all musicians. In a footnote, he made the astonishing broad claim that "musicians, particularly those with high rhythm aptitudes, insist that . . . regardless of tempo, two measures written in  $3/4$  sound the same as one measure written in  $6/8$ " (1976, p. 33). Yet, because there are those who suggest that they audiate  $3/4$  as three tempo beats divided in two, Gordon did include patterns in  $3/4$ . He also included  $9/8$ , though patterns in both  $3/4$  and  $9/8$  appear to be included as a concession.

In 1977 and again in 1980, Gordon took the next logical step and completely did away with  $3/4$  and  $9/8$  in the classification and taxonomy. "It should be understood that music written in  $3/4$  and  $9/8$  is audiated as usual triple meter. That is, one measure of  $3/4$  is audiated as half a measure of  $6/8$ , and one measure of  $9/8$  is audiated as half a measure of  $6/8$ " (1977, p. 54). Gordon's classification and taxonomy were brought into line with his basic premise that "regardless of how rhythm is written, we audiate in pairs of tempo beats" (1977, p. 55).

In an attempt to support his theory, Gordon referred in his works to historical studies on rhythm, but does not give them sufficient prominence to be in the body of the text. He merely cites in footnotes entire works by such diverse (in time as well as in view) authorities as Willi Apel, Thomas Morely, Robert Donnington, Paul Creston, Curt Sachs, Grosvenor Cooper, and Leonard Meyer. And by no stretch of the imagination can tempo beats, meter beats, (or macro and micro beats) and melodic rhythm be thought of in terms of *modus*, *tempus*, and *prolation*, respectively (1976, p. 31). It is possible that Gordon chooses the term "tempo beat" with reference to "tempus" in mensural notation,

which similarly distinguishes duple (*tempus imperfectum*) and triple (*tempus perfectum*) divisions.

In the interest of developing a practical taxonomy, Gordon posits a unique system of meter classification in place of a more traditional one, but it cannot be said to serve any better the needs of either music theorists or educators. It is easy enough to be critical of any given theory of rhythm, none of which has done justice to the complexities of that mysterious part of created reality called "rhythm." As a music education researcher, Gordon is to be commended for inventing terms, which many other theorists do as well.

Gordon is also to be commended as a music education researcher for tackling the difficult area of rhythm theory, especially since so few music theorists have concerned themselves with music learning and teaching theories. One reason for that neglect is that music education research has long been dominated by the behaviorist approach, which concentrates on objectifying components of sonic events rather than musical relationships. The vast majority of research in perception does not address itself to music at all, since the variables in a contextual situation are often too great for statistical analysis.

### Selected Texts by Gordon

#### *Psychology of Music Teaching*

This book (1971) was written 20 years ago, and Gordon's thinking on many of the topics has changed; at least one obtains that impression from the numerous versions of his publications in pattern sequencing. *Psychology of Music Teaching* (1971) was part of a valuable series of paperbacks intended to supplement *Foundations and Principles of Music Education* (1972) by Leonhard and House. Leonhard realized that the treatment of specialized topics including philosophy, research, evaluation, and psychology was inadequate for the advanced student. He assumed the editorship of *Contemporary Perspectives in Music Education* and commissioned the respective authors.

Gordon compares his book with Mursell's *The Psychology of Music School Teaching*

(1931), but Gordon's contribution is not as comprehensive. Gordon focused his thinking on those areas in which he had conducted his own research, and the text is more a reflection of Gordon's ideas rather than more general ideas about the psychology of music teaching. Gordon cites the work of his students as well as his own in the more favorable comparisons. Although Gordon holds the position of Carl E. Seashore Professor of Research in Music Education at Temple University, Seashore's ideas on music aptitude receive more than a few pejorative comments from Gordon because of Seashore's elements approach to music aptitude. Gordon's own aptitude test appears to be more musical than Seashore's, but no definitive study indicates that Gordon's test is a more valid predictor of musical talent. The authors of this article make note of this interesting comparison in that Gordon's teaching ideas are among the more "elemental" in 1991, focusing on the mastery of rather small patterns.

In this book, Gordon suggests that musical aptitude is a product of both innate potential and musical exposure. In 1991, he holds to this basic premise for measurement purposes, but he now believes that a student has his or her maximum aptitude at the time of birth. Musical exposure acts to mitigate the negative influences on musical aptitude that occur in the environment after birth. His original belief was that the main dimensions of music aptitude were rhythmic, tonal, and aesthetic-interpretive (Gordon, 1971, p. 7). This belief apparently has not changed, although the emphasis is shifting away from aesthetic interpretation. His second chapter provides a good review of earlier music aptitude batteries and contains an interesting chart outlining possible components of music aptitude (1971, pp. 26-27). In this chapter, Gordon recognizes the contributions of Seashore and Wing to his own thinking, but most of the chapter is devoted to describing the MAP.

Part II of the text is entitled "Musical Achievement," but Gordon makes no effort to consider many of the extant objectives of music programs. He applies learning theories to music education and selects

Robert M. Gagné as the first psychologist to discuss. Gordon argues that the general purpose of music education is to teach for musical understanding, and he cannot resist informing the reader how to teach. Gordon is a fine teacher; he continually reminds the reader of sound before sign, quoting Lowell Mason and others.

At this time (1971) in Gordon's thinking, music imagery constituted music aptitude (1971, p. 62). Today Gordon's definition of imagery has become more precise. He suggests that students infer musical meaning from musical song by remembering, organizing, and conceptualizing what they perceive; thus, the general purpose of music education should be to teach students to understand the music they hear (1981, p. 63).

### ***A Music Learning Theory for Newborn and Young Children***

Gordon would want reviewers to recognize his earlier contributions to music education, which we have attempted to do, but Gordon modifies his definitions and ideas with some regularity, and knowing his present position on music teaching and learning is difficult. Gordon's most recent book (1990), however, likely reflects his present thinking on the major issues. Gordon stretches the definition of early childhood education in this book in order to present the philosophy that underlies his present research. Gordon is convinced that the first years of a child's life hold the key to music learning. The importance of an aptitude test in the instructional process continues to dominate his thinking.

An interesting point is Gordon's distinction between remedial instruction and compensatory instruction; he suggests that remedial instruction is really not possible. "What was not developed in early life cannot be developed in later life to the extent that it could have been developed in early life" (1990, p. 1). In compensatory instruction, according to Gordon, a teacher assists a child in progressing beyond whatever is brought to formal schooling.

Gordon's use of language is always skillful in arguing for a role for instruction. He maintains, however, that a student cannot progress beyond the musical limits established by his or her inherited ability. He also

posits "all things being equal, a child who begins to learn at a later age will never learn as much as a child who begins to learn at an early age" (1990, p. 1). If Gordon is correct, two things are important: We should not eliminate music from the elementary schools before eliminating it at the secondary and middle schools, and instrumental music teachers who delay the start of instruction until middle school should be warned of the consequences.

Gordon clarifies the difference between guidance and instruction, an idea not developed extensively in his earlier writing. Guidance can be formal and informal and is present in the home, during the critical period in a child's life. Instruction, which has allotted time periods, is characteristic of school instruction and thus has been the focus of music education. Gordon sees that this is changing, or should be changed.

Although music is a literature and not a language, children learn music in much the same way that they learn a language (1990, p. 4). In this sense there are similarities but also significant differences with Sloboda (1988), who draws a number of parallels between language using the structure of the two disciplines as the basis of comparison with Chomsky and Schenker as the primary architects.

Learning stages are crucial to Gordon's learning theory, especially the tonal-babble and the rhythm-babble stages. He believes that children must develop listening and speaking vocabularies in music or they cannot understand and relate to music of their culture. Musical guidance must be provided before a child is 18 months old, for after that time the child becomes preoccupied with language. Gordon further suggests (1990, p. 10) that the quality of the expectant mother's musical environment is important in establishing the talent of the unborn child. This concern is reflected in his statement "the level of music aptitude with which the child is born begins to decrease shortly after birth." A child's music aptitude continues to decrease until his music environment becomes appropriate. Therefore, it cannot rise above birth level (1990, p. 9). The environment of the child is important until the child's musical

aptitude stabilizes at or before the age of 9.

Thus, Gordon focuses on teaching primarily to minimize the adverse effects upon one's inborn music aptitude. "A child cannot be taught the potential to audiate but can be taught how to develop her potential and taught what to audiate. That is music achievement" (1990, p. 18). Developing one's potential, then, is the damage-control program that is initiated after birth; stopping a decline in ability requires a different approach to education than has been the tradition among music education teachers and researchers.

During these early years, children develop through sequential stages that result in preparatory audiation and then audiation. These stages are called acculturation, imitation, and assimilation. The applicability of the stages to all children is reflected in Gordon's concern that high-aptitude students should not be allowed to progress through the stages at a faster rate than students with lower aptitude (1990, p. 31).

In the acculturation stage, there are three substages, the earliest of which extends to 18 months, and then two stages that are not age-dependent but involve (1) musical babble not in response to the environment and (2) musical babble in response to the environment. In his comparison with language development, Gordon differentiates between music babble and language babble and suggests that the guidance provided the student must be adequate if the child is to "safely emerge" from the music-babble stage (1990, p. 14). Guidance is informal, both structured and unstructured, but babble is so critical that an individual with high talent could be impaired if allowed to leave the musical babble stage too soon (1990, p. 130).

The imitation stage has three substages during which the student learns about sameness and differences, imitates with precision, and has the opportunity to hear varied tonal and rhythm music. In the assimilation stage, the student can coordinate his singing of tonal patterns with his muscular movement and with his breathing. Preparatory audiation is important, as the student must be able to sing tonal patterns and to chant rhythm patterns before being

able to sing a song (1990, p. 6).

Gordon continues to make statements based upon his personal experience or perception as opposed to his research finding, a point that was made about comments in his test manuals. For example, Gordon suggests that it is crucial for students to perform alone (1990, p. 7). Supporting this, he again compares learning music to learning a language, arguing that if students were to speak only in groups, they would not learn to use the language. For Gordon, language and music have a biological basis; thought and audiation have a psychological basis (1990, p. 19).

### **Gordon versus Reimer: A Contrast of Ideas**

The stimulation in a profession comes with wrestling with ideas, with the process of sorting out the claims made by the best minds in the profession. In making comparisons, we find that Bennett Reimer promotes an opposing program for music instruction. He posits that the major purpose of education in the arts is to help individuals gain access to the experiences of feeling and suggests that education in the arts is the education of feeling (Reimer, 1989, p. 53). Reimer denies the relationship of music and language: "What takes place between the artists and the perceiver is not communication but sharing" (1989, p. 67). In fact, much of Reimer's (1989) book is devoted to describing how music differs from language.

If a purpose of music education is to develop the ability of students to perceive the embodied, expressive qualities of objects and to react to the intrinsic significance of the qualities in those objects (p. 106), Reimer's educational efforts would concentrate on perceiving the characteristics of sounds which make them expressive. Reimer, but not Gordon, willingly accepts Meyer's ideas of how tonal events arouse tendencies, cause expectations, and produce various kinds and degrees of resolutions and satisfactions. Whether audiation is a step toward perceiving this tension and response, toward regarding music as an expressive form rather than as a symbol, is doubtful. Audiation may not interfere and could be

helpful, but audiation would become a minor objective. For Reimer, language is a powerful tool for increasing aesthetic sensitivity when it is devoted to the refinement of aesthetic perception in context (Reimer, 1989, p. 117). For him, musical tastes can be improved, music enjoyment can be refined, and musical significance made more available to all individuals (1989, p. 134).

Reimer states (1989), "An argument has been made that programs should stem not from a philosophy, but from a psychological theory of how children learn or from learning theory (a term since abandoned by psychologists)" (1989, p. 149). In suggesting that learning theories cannot provide the structure for the learning process, Reimer suggests that psychology can be applied to anything from music to housecleaning, but psychology cannot differentiate between the teaching of cooking and the teaching of music. That is, it cannot provide the answer of what should be taught to all students and what, within the subject (music), should be chosen to be taught—only philosophy can do that because these are questions of value (1989, p. 150).

For the cognitive philosopher, a curriculum must have a plan to select essential subject matter content, an organization of that content appropriate to both the subject (music) and to the cognitive capacities of the learner. "A curriculum to be valid must be inclusive of all possible musical experiences and modes of involvement and developmental patterns" (Reimer, 1989, p. 150).

Reimer states that those who follow learning theory have led us to programs so narrow in goals and behaviors and involvements as to set music apart from any other subject the schools offer. Reimer, in arguing for the basic equality of music and the arts, argues that the arts as a curricular subject do not differ from other school subjects.

Reimer also suggests that curricula based on learning theory become "so technique-skill-proficiency oriented" that there is no basis for the ongoing development of cognitions, including skills, that any bona fide educational curriculum must provide." He states: "Behaviorism as a theory is largely defunct, but some of its practices hold over."

## Colwell and Abrahams: Edwin Gordon's Contribution

For him, "cognitive psychology is changing, often radically, how teaching and learning are being approached" (1989, p. 154).

In another major departure from the writings of Gordon, Reimer suggests that only expressive music be used and that any sequencing of learning should be authentic not only to the students but to the subject as well, and that ways to demonstrate that learning is taking place be employed. Any sequence of learning must be adaptable to schooling in the present culture. Sequencing, for Reimer (1989), is accomplished though the study of significant musical characteristics, including listening, singing, playing, moving, composing, improvising, conceptualizing, analyzing, and evaluating (1989, p. 160). He considers as a serious error the systematization of any phase of the curriculum based on a limited set of musical behaviors where skills become the end rather than a useful means (p. 159).

Reimer also suggests that it is time for us to adapt ourselves to post-Piagetian child development generally and specifically in music (1989, pp. 155-56). Our understanding is that Reimer believes that music education must recognize the most recent research in human development. Literacy for Reimer means "well-educated," and it is imagination in a work of art that captures our feelings.

Reimer argues that eclecticism has gone too far in music education. Too many music educators believe that all methods can be assembled, all philosophies fit together, and all curricula made to work in some situation at some time. Yet separate little pieces do not constitute music, nor do separate little ideas about teaching and learning satisfy the basis for teaching and learning. Reimer calls this eclecticism a collection of inadequacies (1989, p. 155). He conceives of music as a part of general education. Gordon's position is less clear on this point, but he does emphasize providing specialized instruction for the more talented students.

## References

Brink, Emily (1983). "A Look at Edwin E. Gordon's Theories," *Bulletin of the Council for Research in Music Education*, Volume 75, pp. 2-14.

- Colwell, Richard J. (1972) Review of A Three-Year Longitudinal Predictive Validity Study of the *Musical Aptitude Profile*. *The Seventh Mental Measurements Yearbook*. Highland Plains, NJ: Gryphon Press, pp. B246-B247.
- Colwell, Richard J. (1986) Review of *Primary Measures of Music Audiation*. *Test Critiques* Vol. V. Kansas City, MO: Test Corporation of America, pp. 351-357.
- Cooper, G. and Meyer, L. B. (1960) *The rhythmic structure of music*. Chicago: University of Chicago Press.
- Gardner, Eric, et. al. (1982) *The Stanford achievement test*. New York: The Psychological Corp.
- Gordon, E. E. (1967) *A three-year study of the Musical Aptitude Profile*. Iowa City: University of Iowa Press.
- Gordon, E. E. (1970). *Iowa tests of music literacy*. Iowa City: Bureau of Educational Research and Services, University of Iowa.
- Gordon, E. E. (1971) *The psychology of music teaching*. Englewood Cliffs, NJ: Prentice-Hall.
- Gordon, E. E. (1976) *Tonal and rhythm patterns: An objective analysis*. (1976) Albany: State University of New York Press.
- Gordon, E. E. (1977) *Learning sequences and patterns in music* (rev.). Chicago: G.I.A. Publications.
- Gordon, E. E. (1979) *Primary measures of music audiation*. Chicago: G.I.A. Publications.
- Gordon, E. E. (1980) *Learning sequences in music: Skill, content, and patterns*. Chicago: G.I.A. Publications.
- Gordon, E. E. (1981) *The manifestation of developmental music aptitude in the audiation of "same" and "different" as sounds in music*. Chicago: G.I.A. Publications.
- Gordon, E. E. (1982) *Manual for the Intermediate Measures of Music Audiation*. Chicago: G.I.A. Publications.
- Gordon, E. E. (1984) *Learning sequences in music: Skill, content, and patterns*. Chicago: G.I.A. Publications.
- Gordon, E. E. (1986) *Instrument timbre preference test*. Chicago: G.I.A. Publications.
- Gordon, E. E. and Woods, D. G. (1984) *Jump right in: The music curriculum*. Chicago: G.I.A. Publications.
- Lorge, Irving, and Thorndike, Robert, et al. (1954-1966) *Lorge-Thorndike intelligence tests*. Chicago: Riverside Publishing Co.
- Mursell, J. (1931) *The psychology of music teaching*. New York: Silver Burdett.
- Prescott, George A. et al. (1982) *Metropolitan achievement test*. New York: The Psychological Corp.
- Reimer, Bennett. (1989) *A philosophy of music education* (2nd ed.) Englewood Cliffs, NJ: Prentice-Hall.
- Seashore, Carl E. (1960) *Measures of musical*

*talents*. New York: The Psychological Corp  
Sloboda, John A. (ed.) (1988) *The generative  
process in music: The psychology of perform-  
ance, improvisation, and composition*.  
Oxford, England: Clarendon Press.  
Wehner, Walter. (1985) Review of *Primary*

*Measures of Music Audiation. The Ninth  
Mental Measurements Yearbook*. Lincoln, NE:  
University of Nebraska Press, pp. 988-990.

## ***Announcing***

Volume II, Number 3

Fall, 1991



# *The Quarterly*

## Philosophy in Music Education

**Don't miss this important issue!**

"Education is a dynamic enterprise. It is not a static entity. . . . Like music, education may become formalized, even fossilized, as civilizations develop, especially if its instructional methods, curricula, and administration come to be set in stone, regarded as ends rather than means to other ends, and thus cease to be relevant to the life of the surrounding society."

--Estelle Jorgensen

In this issue, outstanding authors will address the issues facing philosophy in music education, discuss the impact of music education philosophy on the field, and debate varying perspectives of the value and definition of music education. Look for articles by these leading music education philosophers:

Bennett Reimer  
Estelle Jorgensen  
David Elliott  
Franz Roehmann

Wayne Bowman  
Jim Johnson  
J. Terry Gates  
Gerald Kneiter

*This issue slated for publication in September, 1991. Don't miss it!*