16

MUSICAL ABILITY

Rosamund Shutery-Dyson

East Grinstead
West Sussex, England

I. CONCEPTS OF MUSICAL ABILITY

The term musical ability is "the broadest and safest" in that it suggests the power to act but indicates "nothing about the heritability or congenitalness of inferred potentiality" (Farnsworth, 1969, p. 151). Such a generic term as musical ability can include a wide range of listening, performing, analyzing, and creating tasks (Boyle, 1992). A broad distinction should, where possible, be drawn between aptitude and achievement. Music aptitude is the term used to indicate potential for learning music, particularly for developing musical skills. However, all aptitude tests are to some extent achievement tests, just as all achievement tests necessarily reflect the initial aptitude the individual can bring to the learning situation. However, accomplishment depends not only on aptitude but also on the teaching received and the child's interest in music and willingness to learn. High achievement in music requires rigorous and intensive training (Sloboda, 1994). Willingness to engage in long hours of practice is associated with certain personality characteristics of musicians (Kemp, 1996).

A useful concept of "developmental" musical aptitude has been put forward by Gordon (1979, p. 4). In his view, the level of musical aptitude a child is born with cannot be raised. However, up to the age of 9, much can be done to ensure that this innate level is in fact realized. After the age of 9, musical aptitude stabilizes. The relative standing of students on the Musical Aptitude Profile (MAP, see Section IIA) does not increase with practice and training. This tends to be true also for the Seashore, Drake, Wing and Bentley tests (see Section II, A), which were intended as group tests for children more than 8 or 9 years old.

A really extensive investigation of representative samples of twins, along the lines of the Minnesota Twin Study, might throw light on how much of musical aptitude is innate (Shuter-Dyson, 1994). Pending such a study, a useful analogy
may be drawn with intelligence. As Vernon (1968) noted, we need to recognize the existence of genetic differences between individuals in their capacities for building up neural and mental schemata, what Hebb (1949) called "Intelligence A." The psychologist, however, can only observe the effectiveness of present behavior or thinking through the interaction of the genes and the stimulation offered by the environment. (i.e., Intelligence B). This is culturally conditioned by the environment provided by the ethnic group into which the child is born, especially by the child's home and his or her leisure pursuits. The effects of physiological conditions before and after birth are also important. The results of intelligence tests, Intelligence C, are of considerable predictive value in so far as they sample useful mental skills.

It is well to remember that the original aim of testing was to provide objective means of assessing abilities, so that justice could be done to the individual. Whereas in the case of music, stress tended to be laid on identifying talented children who might benefit from instrumental lessons, emphasis is now placed on the judicious use of tests as an aid to the teacher in assessing the strengths and weaknesses of the student (Gordon, 1979, p. 6). High scores are indicative of a promising level of talent; low ones may often be due to misunderstanding of instructions or some upsetting circumstance and should be treated with caution. In any case, the results of testing should be supplemented by information from other sources.

The term *musicality* was used by Révész (1953) to denote "the ability to enjoy music aesthetically." The ultimate criterion is the depth to which a person, listening and comprehending, can penetrate the artistic structure of a composition. As Dowling and Harwood (1986) point out, music in industrialized societies comes to depend on high levels of expertise for its performance. The rest of the people become "spectators." Those with initially low levels of ability are often discouraged from active participation in music.

A different concept of musicality implies that all humans qua human can make music, and "average musical ability" is as universal as average linguistic competence (Blacking, 1971). For example, the Venda in Africa assume that everyone is capable of music making; even deaf persons can dance. However, they recognize that some persons perform better than others. Exceptional musical ability is indeed expected of children born into certain families. Only a few may become exceptional musicians; if they do, what is considered to set them apart is that they worked harder at music.

Despite the difficulties of research with very young children, much evidence is accumulating on the inherent music-making potential of humans (see Trehub, 1993). Two-month-old infants were able to match the pitch, intensity, and melodic contour of their mothers' songs, and at 4 months old, could match the rhythmic structure (Papousek, 1982). Very young infants are able to abstract the pitch contour of a sequence of tones (Trehub, 1985; Trehub, Bull, & Thorpe, 1984). They are sensitive to simple rhythmic change (Chang & Trehub, 1977) and to the pitch and durational contours that cue phrase endings (Krumhansl & Jusczyk, 1990). Infants 7 to 10 months old successfully discriminated a semitone change in a well-
structured Western melody but not in a poorly structured Western melody or a
non-Western melody (Trehub, Thorpe, & Trainor, 1990). Kessen, Levine, and
Wendrich (1979) found that before the middle of the first year of life, babies can
learn to imitate pitches and appear to enjoy doing so; they suggested that the infant
brings a congenital readiness to respond to pitched tones. However, most infants
may lose the ability with the onset of learning language and lack of support from
the world around them. Three years later, Wendrich (1981) followed up 9 of the 23
babies who had learned to vocalize the pitches. Seven had lost the ability.

Attentive listening to music and singing before speaking are often assumed to
characterize children with conspicuous musical talent (Shuter-Dyson, 1985).
Many such children are born into families of musicians, but in the case of others,
the families display only modest musical abilities. Howe, Davidson, Moore, and
Sloboda (1995) questioned the parents of 257 children for early signs of musical
ability. All the children had studied a musical instrument but differed in their
achievements. Singing by the child at an earlier age was the only sign observed
that distinguished those children who later succeeded in being accepted by a
highly selective music school. This does not, however, rule out a higher incidence
of early signs of musicality occurring in musicians who as adults reach the highest
peaks of excellence.

II. STUDIES OF MUSICAL ABILITIES

A. METHODS OF STUDY

In discussing studies of musical abilities, we shall look both at the contribution
made by tests and at some of the research that has aimed to understand the cogni-
tive requirements of music and has thereby provided valuable insights into the
abilities involved.

The best of the published tests, in spite of various imperfections, have proved
successful in helping to identify the relative status of individuals to an extent that
makes the tests useful in music education. Detailed information on tests is readily
available elsewhere (George, 1980; Shuter-Dyson & Gabriel, 1981). It is as re-
search tools that tests are to be considered in this paper. In assessing the meaning-
fulness of studies based on testing, we must bear in mind the reliability and valid-
ity of the tests used and their suitability for the groups studied.

Tests naturally reflect differences in their authors' concepts of musical ability.
Seashore (1938, p. 2) believed that the sensory capacities measured in his Mea-
sures of Musical Talents were basic to musical aptitude, could be sharply defined,
and could be absent or present in the individual in varying degrees. In fact, corre-
lation studies of the Measures rarely produce zero correlations (see Shuter-Dyson

Authors of later tests sought to produce material that would seem more relevant
to functional music activities. For example, Drake (1933a, 1933b, 1957) devised a
test of musical memory, where the testee has to remember the original version of a
melody while hearing a series of versions that may vary from the model because of
change of key, time, or notes. Wing (1948) tried out many possible tests; 25 of
the most promising were eventually reduced to 7: 3 of ear acuity (chord analysis,
pitch, and melodic memory) and 4 of appreciation (rhythm, harmony, intensity,
and phrasing). They became the Wing Standardised Tests of Musical Intelligence
(Wing, 1981). Bentley (1966) devised the first battery intended primarily for chil-
dren 7 or 8 to 14 years old, the Measures of Musical Abilities, testing pitch dis-
crimination, tonal memory, chord analysis, and rhythmic memory. Mills (1988)
has produced a test in which the testee is asked to assess which of two tones is
higher in pitch; if no difference is noticed, a “zero” answer is marked.

The MAP (Gordon, 1965, 1988) consists of three parts: tonal (melody and har-
mony), rhythm (tempo and meter), and musical sensitivity (phrasing, balance, and
style). Because rhythm and melody interact in an inseparable way, Gordon consid-
ered it appropriate to test tempo and meter in melodic contexts. His later, shorter
tests, the Primary Measures of Music Audiation (PMMA; Gordon, 1979), the In-
termediate Measures of Musical Audiation (IMMA; Gordon, 1982), and the Ad-
vanced Measures of Musical Audiation (AMMA; Gordon, 1989a) have separate
tonal and rhythm subtests. The three batteries cover age ranges from kindergarten
to adult. Audie (Gordon, 1989b) is intended for 3- and 4-year-old children.

Umemoto, Mikumo, and Murase (1989) have produced a test of pitch devia-
tion. The testee has to find which tone deviates by 50 Hz up or down from the
correct pitch of the diatonic scale with equal temperament. Tonal and atonal se-
dquences four, five, and six tones long are used. Unlike the isolated tones used by
Seashore, the judgments take place within the framework of the scale.

Tests of vocal or instrumental performance have lagged behind perceptual
ones. The Watkins-Farnum Performance Scale (Watkins & Farnum, 1954) con-
sisted largely of sight-reading exercises. McPherson (1993/1994, 1995) has de-
vised tests of ability to (a) play from memory; (b) play by ear and (c) improvise.
Scores were compared with sight-reading (as measured by the Watkins-Farnum
Scale) and with results from the Australian Music Examination Board.

A promising development for the future may be adaptive tests presented by
computer. Correct responses are followed by more difficult ones, incorrect respon-
ses, by easier ones, enabling the testee to answer only those items that are of ap-
propriate difficulty. From a pool of more than 300 tonal memory items, Vispoel
(1993; Vispoel & Coffman, 1992) selected 180 that could be presented by micro-
computer. The synthesizer-produced tonal and atonal melodies were from four to
nine notes in length, the examinee having to judge which if any tone had been
changed on the second playing. The results were compared with results of Sea-
shore’s, Wing’s and Drake’s memory tests and with results of Gordon’s AMMA.
Reliability and validity levels higher than with fixed-order tests were obtained in
much shorter time. This approach might enable the reliability and validity of a
longer test to be retained with reduced length or the reliability of batteries with
short subtests to be improved. So far Vispoel has used material based on memory
tests, but the procedure could be extended to other musical abilities. He has also experimented with self-adapted testing, where the examinee is allowed to choose the difficulty of the items. This procedure produced some loss of testing efficiency, but less anxiety among the examinees (Vispoel & Coffman, 1994).

B. FUNDAMENTAL ABILITIES

There now seems to be agreement on the importance of perception and cognition of patterns and structures. Before looking at studies of components of musical ability, such as tonal, rhythmic, and kinesthetic abilities, we shall first discuss the work of Gordon, of Karma, and of Serafine in their attempts to throw light on the more general ability to deal with patterns and structures and the temporal aspect of music.

Gordon (1979) uses the term audiation to mean "the ability to give meaning to what one hears." Gordon (1990) succinctly described five stages in the process of audiation: The first stage is when one perceives the sound. The second is when one begins to give meaning to sound through tonal and rhythmic patterns within a context of tonality and meter. These first two stages represent musical aptitude. At the third stage, one asks "What have I just heard?" and begins to find meaningful answers. The fourth stage is marked by the question "Where have I heard these patterns and sounds before?" and in the last stage, one begins to predict what one will hear next. These last three stages are achievement.

Audiation is thus the essential cognitive function that not only enables persons to give meaning to music while listening, but also enables them to bring order and meaning to music read, or written from dictation, recalled from the past or improvised. (For a discussion of Gordon's contribution to research and music learning, see The Quarterly, 1991).

Karma (1985) defined musical aptitude as "the ability to structure acoustic material." He believes that the sense of tonality, of rhythm, and of harmony can be seen as culture-specific reflections of a general structuring ability. Because it is difficult to remember unstructured material, musical memory can be seen as the consequence of conceiving the structure of the music heard. Karma has developed a test in which a short sequence of tones is repeated three times. Only one factor—pitch, intensity and length—is varied within an item. This theme must be compared with an "answer" that follows after a pause and judged as "same" or "different." Testees tend to adopt two strategies: anticipating and recognizing. Items where the structuring conforms to strong gestalts, where, for example, a stressed tone occurs at the beginning of a subgroup, are easy. Ability to structure against strong gestalts requires greater skill, as do items where the beginnings and the ends of the subgroups are similar, as these require changing expectations.

Karma (1986) reported that "changing expectations" and "structuring against strong gestalts" are significant predictors of the ability to solve difficult items, both among musically select candidates for entrance to music school and among
musically unselected children. Sensory discrimination was incorporated into the test by making the pitch, length and intensity differences relatively small in some of the items.

Karma (1991) has also experimented with tests to measure as purely as possible the ability to perceive pitch differences and their direction, comparing the results with a test that requires verbal labeling. The pitch-differentiation test consisted of 24 tone pairs, with the testee having to say “up” or “down.” The intervals range from a perfect fifth to 13 Hz. In a “same/different” test, an item consisted of two three-, or four-tone patterns that might either be similar (same) or inversions of each other (different). This test appears to be an unusually pure test of pitch-differentiation ability but requires further experimentation. The traditional up/down version does measure pitch differentiation but is unsuitable for use with very young children.

Serafine (1988) stresses the temporality of music. To investigate successive processes, she devised tasks of phrasing (dividing a span of music into two parts), patterning (following a pattern of alternating fragments to tell which would come next), motivic chaining (understanding that motive A followed by B will form a longer phrase AB rather than AX or BZ), and idiomatic construction (discriminating coherent melodies from random sequences of tones). Simultaneous processes were examined by a textural abstraction task (listening to two or three parts and deciding how many parts are being heard), motivic synthesis (will fragments A and B, if played simultaneously sound like A with B rather than A with X or Z with B?), timbre synthesis (are two instruments, first heard separately, being played together, or is the second instrument a different one)?

Serafine also proposes four types of nontemporal process, concerned with the overall, formal properties of music: closure (which fragment is “finished”?), transformation (recognizing a theme as similar when transformed, as opposed to a foil), abstraction (recognizing a melodic or rhythmic pattern when it occurs in a new context), and understanding the hierarchical levels of musical structure.

To trace the development of the abilities required, Serafine tested individually children and adults, 168 in all, with approximately 30 at each age level, 5, 6, 8, 10, and 11 years old and 15 adults. By the age of 10 or 11, most of the processes had been acquired, with the exception of the ability to identify the number of simultaneous parts in a complex texture. The 8-year-old children gave evidence of perceiving hierarchic levels in simple melodies, they did well at perceiving simultaneous combinations of timbres, and performed as well as older children did at discrimination of random melodies, although unlike the older groups, they did no better on melodies constructed in accordance with tonal idiom. It seemed that nontemporal processes developed earlier. Serafine suggested that the child first makes sense of global features and only later is able to give the constant analytic monitoring required by temporal tasks.

Zenatti (1985) reported a factor analysis of the scores of 89 children from 6 to 7 years old on 17 memory tests and confirmed that children whose perception tends to be global form a separate group from children who have more analytical
perception and possess more developed musical discrimination capabilities. Thirty-four children who had received intensive Suzuki training were tested on closure, transformation, and hierarchical levels. Only those who had received very intensive, long-term training that was specifically relevant to certain tasks achieved higher scores than children of the same age. With age being the principal predictor of success, the most potent factors may be general cognitive growth and normal, everyday musical experience.

Although Serafine was concerned with demonstrating the universality of musical cognition, many individual differences were found among her results. It would be interesting to know how the children would have scored on musical aptitude tests. She did include a pitch-discrimination test, in which her subjects were asked whether pairs of prerecorded piano tones were the same or different. Where significant correlations with the other tasks occurred, they were low.

To sum up, all three, Gordon, Karma, and Serafine, are concerned with means of making sense of music in time. Memory for music as such is not the primary interest. Gordon’s concept of audiation is redolent of Seashore’s (1938, p. 6) statement that “tonal imagery is a condition for learning, for retention, for recall, for recognition, and for the anticipation of musical facts.” Both the Gordon and Serafine tasks are affected by age. Karma (1983) claims that scores on his structuring task rise by only 10% between 8-year-olds and adults. His test should be applicable to non-Western musics. Although her approach was not psychometric, Serafine’s varied tasks cover a wide and interesting range of abilities. For her purpose, the tasks had to be accessible to 5-year-olds yet require no major change for older children and adults. With the youngest children, they were presented in a play situation. For a succinct discussion of her theoretical position, see Hargreaves and Zimmerman (1992).

C. TONAL ABILITIES

1. Pitch Perception

Pitch discrimination is widely acknowledged to be an important part of musical perception. Pitch tasks are among the most difficult for children younger than 5 or 6 years old, even when the tones can be presented by a touch screen where the child has only to touch part of a screen to respond (see White, Dale, & Carlsen, 1990).

Quite strong intercorrelations are typically found between tests of fine pitch differences (Seashore and Bentley) and the Wing test, where the pitch change is masked within two piano chords. Results of both the Seashore and the Wing test correlated quite highly with results of the Lundin test, where the listener has to judge whether a second melodic interval moves up or down (see Shuter-Dyson & Gabriel, 1981, pp. 58–59).

Umemoto (1995) carried out a factor analysis of the pitch-deviation test, tests of pitch intervals and absolute pitch, along with scores from the Seashore Measures, the Wing and the Drake tests, the sensitivity subtests of Gordon’s MAP, and
the five subtests of the Sherman-Knight battery (tests where recorded excerpts must be compared with notation). The test of pitch deviation had a high loading of 0.785 in the first factor. The other tests that had loadings greater than 0.3 on this factor were the tests of absolute pitch, the test of intervals, all the Sherman-Knight tests, the phrasing test of the MAP, and tests 1–4 of Wing. The Seashore pitch test formed a separate factor, on which the Wing and Chord Analysis tests had some loadings. Umemoto thus claimed that the sense of pitch deviation has a central position in the structure of musical ability. It also had a loading of 0.448 on a second factor where the memory tests of Seashore, Wing, and Drake also had high loadings. This is not surprising, as memory is involved in the test.

The highest loading on the first factor was 0.822 for the Absolute Pitch test. The value of absolute pitch to musicians has been much discussed. Miyazaki (1992, 1993) believes that the possession of absolute pitch may impede the development of relative pitch, which is much more important in musically meaningful situations. He agrees, however, that many musicians have both. Absolute pitch certainly appears to be a characteristic of many extremely talented performers—and also of “musical savants” (see Section III).

2. Sense of Tonality

Hargreaves and Zimmerman (1992) state that “By the age of 8, children’s melodic perception operates within an increasingly stable tonal system. Now melodic information is stored and processed according to a tonal reference rather than by contour schemes.” The age at which a child develops a sense of tonality depends on the criterion employed and on individual differences.

a. Singing

Davidson, McKernon, and Gardner (1981) found that the 5-year-olds they studied either sang in the same key throughout or changed key suddenly and then stayed within it. Welch (1994) reported evidence in the singing of a sample of 5-year-olds of a “tonal centre” that was located toward the lower part of the singing range. Concomitantly, intervals tended to be reduced in size, especially those further from the tonal center. As might be expected, there is considerable variability among preschool children (Flowers & Dunne-Sousa, 1990). Also, singing ability is an important factor—more than one fourth of preschool children classified as having low singing ability exhibited no evidence of ability to maintain the tonal center of a song (Ramsey, 1983).

b. Recognizing Change of Key

Bartlett and Dowling (1980) found that 5- or 6-year-old children noticed changes of key when the change was to a distant key (i.e., one that introduced several changes in pitches of notes in the scale) but not when the change was to a nearly related key. Boyle and Penticoff (1989) concluded that the average child in Grades 4 and higher can identify tonality change with almost 90% accuracy for familiar melodies and 85% for unfamiliar melodies. Imberty (1969) found that, by
age 7, children could recognize sudden changes of key in the midst of tunes and by age 8, a change of mode.

c. Recognizing the “Tonal Hierarchy”

Cuddy and Wiebe (1979) noted that “the sense of tonality involves the development of a pitch system in which tone relations are specifically defined ... even tones not played in a given melody are represented by their inferred relation to the tonic.” A great deal of research effort has gone into the investigation of the part played by tonality in the cognition of music (see, e.g., Cuddy, 1997; Cuddy, Cohen, & Miller, 1979; Krumhansl, 1990). Two examples concerning children follow: Cuddy and Badertscher (1987) presented the major triad, the major scale, and the diminished triad followed by a probe tone (each of the 12 tones of the chromatic scale) to 53 children, from 6 to 12 years old. The children were asked to rate each of the probes as to whether it sounded “good” or “bad” as an ending to the melodic pattern. The tonic was judged the best ending, especially for the major triad. Speer and Meeks (1985) found that both second and fifth graders chose diatonic over nondiatomic notes, triad notes over other diatonic notes, and the tonic over other triad notes.

d. Better Discrimination with Tonal As Opposed to Atonal Music and Preference for Tonal

Zenatti (1969, 1981) emphasized the importance of acculturation in her studies of the acquisition of tonality. In one of her experiments, children were asked to say which one of three notes in a sequence had been changed in pitch on a second presentation. Five-year-olds performed at chance level with both tonal and atonal melodies, but by age 6 or 7, superior performance on the tonal items began to emerge. In the case of four-tone phrases, the better the discrimination score, the greater was the difference between tonal and atonal perception.

3. Harmony and Polyphony

The acculturation noted above occurs also in harmonic contexts. Children 5 to 7 years old can discriminate a harmonic change more easily when it is presented in consonant, as opposed to dissonant, chords (Zenatti, 1985). Sloboda (1985, pp. 211–213) describes four tests of harmony. Two tests contrasted dissonant with consonant items. Five-year-olds scored very poorly. By age 7, performance had improved significantly, and by age 9, the score was indistinguishable from the scores of adults. In a third test, in one sequence of each pair of items, the chords were played in a “musical” order, leading to a conventional cadence, whereas in the other sequence the same chords were presented in a “scrambled” order. Not until the age of 11 did scores on this test approach those of adults—to detect incorrectness, the ordering of the chords has to be observed. In his fourth test, the items were unaccompanied melodic sequences. One item was a diatonic melody that remained within a single key, the other was a sequence with the same contour but with chromatic notes. This was the most difficult, requiring ability to detect viola-
tions of normal sequential structure. Performance of the tests was not aided by musical training.

Imberty (1969) concluded from his research on cadences that up to the age of 5, a child perceives a tune as finished when the music stops. The effect of a perfect cadence (V to I) is felt by 8-year-olds. By age 10, children understand both perfect and imperfect cadences, but surprise was caused if the melody ended on the third degree of the scale. Serafine (1988) found that 10-year-olds succeeded in recognizing "closure" on harmonic items but scored only at chance with the single-line melodic item.

Although Zenatti (1969, Chapter II) found children between 8 and 10 years old began to be able to recognize a well-known tune presented fugally, even children up to 12 years old had difficulty perceiving the bass part. This result seems in accord with the norms Gordon provides for his tonal (harmony) test, where a judgment has to be made about the lower part. The greatest rise in mean scores takes place between ages 9 and 10, although scores continue to improve until age 17. Wing's chord analysis proved to be effective at separating the good from the very good students at the Eastman School of Music. As noted in Section II,B, Serafine (1988) reported that ability to identify the number of parts in a complex texture was not generally present until adulthood. The examples of the music she quotes on page 147 do indeed suggest that the task is complex. Identifying a pair of disparate timbres seems a simple task, but only about half the 5- or 6-year-old children succeeded (pp.154–156). Even expert musicians found difficulty in identifying concurrent voices in Bach's St. Anne Fugue, although they were more accurate than a nonmusician (Huron, 1989).

C. RHYTHMIC ABILITIES

1. Rhythm/Tonal Studies

Sizable correlations are found between the tonal and rhythmic parts of Gordon's MAP. Gordon (1965) noted that melody undoubtedly influences perception of rhythm. Indeed, as Sloboda (1985, p. 188) comments, not only is rhythm just as important an organizing principle as tonality in music, but both systems are mutually interactive. In much tonal music, knowledge of the tonal structure can help determine the rhythmic structure and vice versa. Sloboda (pp. 52–55) also observed that many of the constraints on melody are dictated by the requirement that the melody be capable of communicating harmonic and rhythmic structure to a listener. An example of the importance of rhythmic structure to memory was provided by an investigation of the immediate recall of a melody by Sloboda and Parker (1985). They asked eight female psychology students to listen to a Russian folk song and to recall it immediately afterward. Six recalls were obtained. Four of the students were musically trained and active performers, the other four were just interested in music. The most fundamental feature recalled was the metrical struc-
ture, preserving the quadruple meter and the two-bar phrase subdivisions. Within the metrical structure, metrical equivalents were substituted about half of the time. The only significant difference between the musically trained women and the other women was the ability to retain harmonic structure.

2. Rhythmic Perception/Performance

Thackray (1969) devised tests of rhythmic perception, rhythmic performance, and rhythmic movement. These were intended for adults; he later adapted the perception and performance tests for use with children (Thackray, 1972). A factor analysis of the correlation coefficients of his perception subtests suggested that the ability to perceive and memorize a rhythmic structure as a whole, and to analyze it consciously, was fundamental to rhythmic perception. Inability to memorize some of the longer items was an important reason for low scores. With adult physical education students, perception and performance had a correlation of 0.63, perception and movement had a correlation of 0.55, and performance and movement had a correlation of 0.59. With child groups, performance tests seemed superior to tests of perception as a means of testing general rhythmic ability.

Hiriartborde and Fraisse (1968) carried out a factor analysis of a series of tests—some predominantly perceptual, some predominantly motor, along with the Seashore and Wing batteries. They concluded that several independent factors are needed to account for the plurality of rhythmic aptitudes that correspond to the many aspects of rhythm. Their factors included perception of rhythmic structures, rhythmic anticipation (synchronization and several Seashore and Wing tests that imply memorizing a pattern in order to compare it with a following one), and a “practo-rhythmic” factor concerned with coordination of limbs in rhythmic movement (see Fraisse, 1982, pp. 176–177).

Thackray (1972) noted that ability to maintain a steady tempo appeared to be a highly specific rhythmic ability. It can be a problem with otherwise promising students. Gordon and Martin (1993/1994) go so far as to say that playing in time is often the most important aspect of a musical performance. Drake’s rhythm tests are concerned with the ability to maintain a steady beat, in the case of Form B against a faster or slower distracting beat. These tests were found to form a separate factor in factorial studies (Tomita, 1986; Umemoto, 1995). Mills (1988) has developed a test on the lines of Drake’s.

Zenatti (1976) found evidence of a stage at about 4 years 8 months, when children showed a considerable improvement in tapping back two-, three-, or four-note rhythms on their second attempt, helped by seeing the experimenter tap the rhythm again. Stambak (1960) used more elaborate patterns that required some measure of temporal structuring. A clear development occurred between 6 and 9 years old. Gérard and Drake (1990) reported from four experiments with children 5 to 8 years old, that whereas 6-year-old children would perceive and reproduce simple time patterns, even 8-year-olds rarely reproduced intensity differences to indicate accentuation.
The experiments of Bamberger (1982) have called attention to the “figural” and the “metric” construing of music. Figural perception groups sounds into meaningful chunks, whereas metric perception focuses on a steady pulse underlying the surface events of melody. Upitis (1987) assessed the figural and metric aspects of rhythm through individual interviews with 72 children, 7 to 12 years old. The children were asked, for example, to describe a clapped rhythm by putting down “whatever you think will help you to remember the piece,” then to “put in some numbers that seem to fit with the marks you have made”; to interpret descriptions of rhythms; to keep time by beating on a drum; and to identify duple or triple beats that would be congruent with unfamiliar melodies. Results showed that children who had had at least 1 year of music lessons were better at keeping time than untrained children. Ability to pick congruent beats was associated with age, but not with training. Upitis’s findings indicated that all the children were able to make sense of rhythm, using either figural or metric techniques, or both. There was a considerable variation in response among individuals. In a later study, accuracy of figural drawings was significantly correlated with ability to join in clapping and to clap back a rhythm (Smith, Cuddy, & Upitis, 1994).

Gromko (1994) found evidence that the notations children invent are a valid indication of their musical understanding, as measured by the PMMA, and their ability to reproduce a short folk tune by singing and playing.

E. KINESTHETIC ABILITIES

1. Instrumental

In a reanalysis of many previous correlation and factorial studies, Whellams (1971) reported a factor he interpreted as “kinesthetic perception,” which seemed to span the tonal memory and rhythmic aspects of music. This factor drew attention to the muscular component in auditory perception that is in line with feedback models of skilled behavior; it has long been noted in connection with music. For example, Mainwaring (1933) found that both children and students of education tended to translate auditory into kinesthetic cues in order to recall tunes they had listened to well enough to answer questions about them.

Gilbert (1979) devised a test of motoric skills. The tasks entailed striking a musical instrument (drums, xylophone, etc.) with a mallet through use of vertical arm and hand motion. She tested 800 children between 3 and 6 years old. A year later she retested 87 of the children. Improvements related to age were found on all the subtests except for motor pattern coordination. The greatest gains were made between the ages of 3 and 4 (Gilbert, 1981). A high correlation was found between results of her test and results of the PMMA (Gilbert, 1982). The effect of psychomotor activities on achievement on the MAP-Rhythm and IMMA-Rhythm tests was investigated by Schmidt and Lewis (1987) with 29 fourth graders. Besides the general music curriculum, the children received eighteen 30-minute class sessions in which the musical concepts of tempo, meter, and rhythm were introduced and
reinforced through activities that required synchronization of body movements while singing or listening to music. The mean scores on the MAP-Tempo, but not on the MAP-Meter or IMMA-Rhythm, were significantly improved on post-test.

Sloboda (1985, pp. 88–89) noted the importance for performing of “motor programming”—the setting up of a sequence of commands to the performing muscles that will reveal in sound the expressive devices selected from the performer’s “dictionary.” The elements in a motor sequence overlap in time—the second finger may already be moving to its position before the first is released. Fluent motor behavior results in specified goals being achieved rather than specified movements being carried out. McPherson (1993/1994) noted that kinesthetic factors such as the organized fingering system of the clarinet seem to influence ability to improvise (see Section II,F).

Baily (1985) emphasized the need to study the way musical patterns may be represented cognitively as patterns of movement rather than as patterns of sound. He believed that the extent to which the creation of musical structures is shaped by sensorimotor mechanisms must also be considered. He cited ethnomusicology studies of African music and his own research in Afghanistan. Overt body movement is a prominent feature in the performance of many kinds of music in Africa, and the conceptual link between music and dance is very close. In villages where everyone participates in music making, movements are very easily observed. J. Davidson (1993) has shown that the bodily movements made by performers contribute to the expressivity of the performance—an element missing for the listener to radio or compact discs.

2. Singing

“Tone deafness” is not just a problem of some boys spoiling class singing. About 10% of 600 university students considered themselves tone deaf and reported a variety of difficulties in melodic perception and memory, vocal production, and auditory imagery (Mawhinney, 1987). They had had few musical activities in early childhood and had unpleasant memories of such musical education as they had received. Kalmus and Fry (1980) investigated “tune deafness” by devising a Distorted Tunes test that required the detection of wrong notes in Western popular and classical music. Familiarity with the actual tunes was not a prerequisite for a good score. From the age of 9, the percentage of children failing according to the adult criterion fell from 40% to the adult value of 4%. Their difficulty reflected a failure to have acquired knowledge of the syntactic rules of the Western tonal system. Scores were only weakly correlated with scores on the Seashore pitch and tonal memory tests. Mawhinney used an adapted form of the Distorted Tunes test with 42 university students. Those who considered themselves to be tone deaf scored 59% correct, compared with the 78% accuracy of those who classified themselves as “above average” as listeners. In a probe-tone test, 6 tone-deaf students out of 41 students produced tonal hierarchy ratings only somewhat weaker than did the other listeners (Cuddy, 1993).
Mitchell (1991) found that three nonsingers were able to perform the Distorted Tunes test at ceiling. Their singing, however, was not only severely out of tune, but they were unaware of this. They first had to learn to recognize a feeling of being in unison with a keyboard or another singer and then to control their voices so as to achieve unison. No subject showed measurable improvement on song performance.

L. Davidson (1993) found that a significant number of musically untrained adults typically compress the range of the third phrase of “Happy Birthday” from an octave to a fifth or sixth, and end the song on a different key, just as young children do (see Section II,C,2,b).

However, from a review of research with children, Goetze, Cooper, and Brown (1990) concluded that children at least can be taught to sing with improved accuracy. Welch (1985) postulated that learning to sing in tune is achieved via a motor response schema. With children who made gross pitch errors and who were given individual training, those who were given knowledge of results gained significantly greater accuracy than those who were not; this was particularly true of the children who were most out-of-tune at the beginning of the training. Visual feedback by means of a microcomputer system was also helpful (Welch, Howard, & Rush, 1989).

F. AESTHETIC ABILITIES

In the Indiana-Oregon Music Discrimination Test (developed by Long, 1965, from the Hevner Landbury test) and the last four tests of the Wing battery, excerpts of original music are paired with a “distorted” version. Wing requires the testee to identify the original version, Long’s testees have also to judge which element—rhythm, harmony, or melody—has been changed. Gordon (1965) composed the items for his sensitivity tests and then asked professional musicians which version they preferred. Only when high consensus (9 out of 10) was reached, and this was confirmed by field trials, were the items incorporated into the final version of the test. Boyle (1982) noted that the correlations among the Wing, Gordon, and Oregon appreciation tests were quite low.

It is interesting that later research is throwing light on the basis for aesthetic judgments. For instance, Repp (1992) obtained tone-onset timing measurements for 28 different performances by famous pianists of a “melodic gesture” from Schumann’s Träumerei. The pianists tended to speed up in the initial part and slow down at the end in a “parabolic” fashion. Listeners were asked to judge a variety of timing patterns (original parabolic, shifted, and nonparabolic) for aesthetic appeal. The parabolic patterns received the highest ratings from musically trained listeners. (Musically untrained listeners did not make consistent judgments). Pianists required to imitate performance of rubato in short musical excerpts were less accurate when the structure of the music and expression conflicted than when they were consistent (Clarke & Baker-Short, 1987). Clarke (1993) found that the more the relationship between structure and expression is disrupted, the more inaccu-
rate and unstable is the attempt at imitation. Moreover, listeners' preferences follow the same pattern as the accuracy/stability measures for the imitation attempts.

Gabrielsson (1982) concluded from his research on rhythm that a balanced combination of the structural, motional, and emotional aspects, adapted to the needs of the specific individual and the actual musical content, may be what is required for artistic performance. This agrees with the results of Sloboda's (1983) investigation into the communication of musical meter. Six pianists of various degrees of experience sight-read a sequence of the same notes but with two different metrical stresses. The more experienced players made greater use of expressive variation and conveyed the differences with more certainty to listeners.

From his experience of teaching music, Sloboda (1985, p. 88) noticed very great differences between even quite young children in the ability to notice expressive variations in the performance of others. Some were capable of immediate and accurate imitation and retention of points demonstrated by the teacher; this seemed to be one of the best predictors of high levels of musical achievement.

Gardner (1973) adopted a different approach: he presented children with pairs of musical extracts and asked them to judge whether the two excerpts came from the same piece of music. He found that the 6-year-olds tended to rate most pairs as "different." The highest absolute scores were made by the 21-year-olds, who were able to base their judgments on variables such as instrumentation and texture. Castell (1982) carried out a similar investigation with 8- and 11-year-olds, including jazz and rock music alongside classical pieces. All the children were more accurate when judging popular music, the 8-year-olds being significantly better than the 11-year-olds; there was no such difference for the classical pairs.

Musical appreciation depends on awareness of basic aural elements. Correlations between the Hevner-Landbury test and Seashore's tonal memory test and the Wing memory, pitch, and harmony tests were quite strong (see Shuter-Dyson, 1982). The "Which element changed?" judgment received the highest factor loading in McLeish's (1950) musical cognition study.

Sloboda (1990) reported the results of an investigation of adults' recollections of their early involvement with music. These seemed to reveal an evolutionary process in the acquisition of musical meaning. The age of 7 seemed to signal the progression to a new awareness—the age when the grasp of tonal syntax is becoming apparent. Sloboda suspected that the syntactic and semantic developments are not unrelated. Swanick (1973) also concluded that much cognitive activity is involved in aesthetic response to music and that the intensity and quality of any emotional experience depends on this activity; ability to make predictions as to what may follow so that deviations arouse excitement is central to the process of understanding music.

G. CREATIVE ABILITIES

Webster (1988) noted that first among the "enabling skills" needed for creative work in music are the "necessary collection of 'musical aptitudes.'" The "conver-
"Innovative" skills include ability to recognize rhythmic and tonal patterns and musical syntax. "Divergent" skills include musical extensiveness (how many ideas are generated), flexibility (ease of moving freely from extremes of fast/slow, soft/loud, high/low), and originality. Other abilities required are conceptual understanding, craftsmanship, and aesthetic sensitivity.

Vaughan (1977) devised a measure of musical creativity on the lines of the Torrance Tests of Creative Thinking. For example, a rhythmic or melodic pattern was presented and the child was encouraged to improvise an answer; or a basic outline or ostinato was set up for the child to make up a pattern to go with it. Wang’s Measures of Creativity in Sound and Music (Wang, 1985) for children ages 3 through 6, require the child to produce examples of a steady beat, imitate events described by the teacher, improvise ostinatos on a bass xylophone, and move to music. Responses are scored for fluency and imagination. Webster also has experimented with measures of creative thinking in music. Version II (Webster, 1983) consists of 10 activities including improvisation tasks for children ages 6 to 10. For example, they are asked to create call-and-response patterns on temple blocks and encouraged to create a composition that uses all the instruments available and has a beginning, a middle, and an end. The Measures are scored for originality, extensiveness, and flexibility, as well as musical syntax.

Such tests of creativity can be reliably scored and the internal consistency of the Webster and the Wang tests are generally high (Baltzer, 1988, 1990; Webster, Yale, & Haefner, 1988). Musical creativity factors seem to be discrete from those assessed by musical aptitude. Swanner (1985) found that results with the Webster test were not related to the PMMA, nor to teachers’ and parents’ ratings of the creativity of 69 third-grade children. Personality traits of imagination, curiosity, and anxiety accounted for 29% of the variance in the Webster test composite score.

Kratus (1989, 1991) believed that how children compose is as important as what they compose. For his 1994 study, he asked 40 9-year-old children to spend 30 minutes making up “a brand new song” on an electric piano. The process of composition—exploration, development, repetition and silence—was analyzed, as well as the product. IMMA test scores were negatively correlated with exploration, indicating that the greater the ability to audiate, the less the need to explore. Audiation was positively connected with development of ideas and silence—children who can hear music inwardly can compose without the sound being physically present. Audiation was also related to tonal and metric cohesiveness and the use of developed, rather than repeated, patterns.

McPherson (1993/1994; 1995/1996) also emphasized the importance of being able to “think in sound” for musical improvisation. He investigated the improvisation skill of 101 high school students of clarinet and of trumpet. His test required the student to provide an answering phrase, improvise “an interesting” melody on a given rhythm and to a recorded piano accompaniment, improvise on motifs, and to produce a “freely conceived” composition. High correlations were found be-
tween ability to improvise and (a) ability to play by ear and (b) ability to sight-read. Performance proficiency on an instrument and ability to improvise become increasingly intertwined in more experienced performers. For a discussion of the processes of composition and improvisation at advanced levels, see Sloboda (1985, Chapter 4).

III. MUSIC ABILITY AND OTHER ABILITIES

Correlations between intelligence scores and musical ability tests are mostly found to be positive, but low—"typically" around 0.30 (Shuter-Dyson & Gabriel, 1981). Wing (1954) observed that there was usually good agreement between low intelligence and low Wing scores, but a high IQ might be accompanied by a low musical ability score.

In the case of those tests of achievement in music that are largely aurally based, correlations with intellectual tests are also low. On the other hand, correlations between intelligence tests and grades on, for example, theory of harmony and the history of music are often much higher.

Because correlations are likely to be depressed by atypical cases, Sergeant and Thatcher (1974) preferred to use analysis of variance. They investigated the interrelations among intelligence, ratings of sociocultural and socioeconomic characteristics of the family, and scores on a rhythmic and a melodic task. Highly significant relationships were found among all the variables. Similar results were reported by Phillips (1976).

Huntsinger and Jose (1991) used four short-term memory tasks (digit recall, tone recall, digit recognition, tone recognition) with children ages 6 to 10. Moderate to strong correlations were obtained on all four tasks, and only one factor emerged from analysis, suggesting that auditory storage in short-term memory may be a unitary phenomenon. However, in a study of 100 adults, some musically trained, others not, Steinke (1992) compared tonal abstraction abilities with non-music abstraction tests, memory for pitch, numbers, and figures. The nonmusic tests loaded on a separate factor from the music tests. Digit Span had a moderate loading on both factors, but the loading on music was much reduced when music training was partialed out.

The concept of musical ability as being specific has been adopted by Gardner (1983), who argued that music should be considered as one of several loosely related multiple intelligences. Among his criteria for an intelligence is the existence of idiots savants and prodigies. The existence of prodigies in music has long been recognized (see, e.g., Shuter-Dyson, 1985). A notable investigation of musical savants has been carried out by Miller (1989). He describes in detail the case of Eddie, whose progress in music, and through music, was followed from the age of 5 to 9½. Miller reviewed previous cases and carried out several carefully designed experiments in which the performance of Eddie and other musical savants was
compared with performance of other child and adult musicians. As mentioned above (see Section II,C,1) an outstanding characteristic was absolute pitch, which enabled some of the musical savants to make confident and rapid judgments not only of the pitch of notes but also of complex chords and to produce results that were remarkably similar to those of competent adult musicians. Like intelligent adult musicians, they were sensitive to the various rules reflecting the harmonic relationships and the structure of musical compositions. However, rhythmic abilities were less in evidence. Many of the musical savants are congenitally blind. This renders them more likely to concentrate on sounds. Again, many suffer from severe language disorders that cause them to be sensitive to speech only as a series of sounds.

Saperston (1986) found highly significant correlations between the language and singing abilities of normal and of delayed young children and of mentally retarded adults.

Barwick, Valentine, West, and Wilding (1989) reported that scores on Bentley’s tonal memory and chord analysis tests were related significantly to reading ability among children between 7 and 10 years old. Rados (1996) found a significant correlation between the Bentley and Wing tests 1–3 and verbal intelligence, especially for musically unselected children. In a study by Hermelin and O’Connor (1980), musically highly talented children were able to access lexical information as fast as more intelligent nonmusical children. This was not due to faster reaction time nor to mere speed of reading. A later study showed that musically talented children of high as well as of more average intelligence performed better with verbal than with nonverbal items (O’Connor & Hermelin, 1983).

The association between musical and mathematical abilities remains unproven (Shuter-Dyson & Gabriel, 1981). Wang and McCaskill (1989) correlated results on two music achievement tests with results on tests of mathematical and visual-spatial skills. Only among girls was a correlation of 0.42 between music and mathematics found. However, spatial abilities appeared to be a significant factor in the prediction of musical achievement. Hassler (1989) also has reported a stable relationship between musical and spatial abilities, both in the case of a longitudinal study of adolescents and among adult musicians.

Rauscher, Shaw, and Ky (1995) postulated that a short-term enhancement of spatial-temporal reasoning produced among college students by listening to Mozart might have a neurophysiological basis. Musical activity may strengthen inherent neural firing patterns in the cortex, which may also be exploited by spatial-reasoning tasks. To test this hypothesis, they provided keyboard training for 34 preschool children for 6 months. Compared with controls, a highly significant improvement was found for the keyboard group on the Object Assembly task, which requires spatial-temporal reasoning, but not on other tests of spatial-logical reasoning (Rauscher, Shaw, Levine, Wright, Dennis, & Newcomb, 1997).

Gardiner, Fox, Knowles, and Jeffrey (1996) reported a study in which four first-grade classes participated in a music and visual-arts curriculum that emphasized
sequenced skill development. Improvements in reading and particularly in mathematics were achieved, perhaps due to a better attitude to learning and to the “stretching” required by the arts skills developing the type of thinking needed for mathematics.

Such results are not incompatible with a lack of definite relationships among older subjects, who are unlikely to have enjoyed enhanced preschool opportunities for music. Techniques such as electroencephalographic mapping and positron emission tomography scans may eventually throw light on the processing of music in the brain and the relationship of music to other abilities.

IV. CONCLUSIONS

There seems to be ample evidence that music is as natural for humans as is language. It is also apparent that musical abilities blossom in a social climate where music is valued and enjoyed.

A musical background very early in life is likely to be most effective in helping individuals to fulfill whatever aptitudes they happen to have been born with, as well as revealing special gifts. The encouragement of excellence must always be an important aim. Manturzewska (1990) has shown how crucially professional musicians need appropriate support at different stages throughout their lives.

L. Davidson (1994) notes that instrumental training in itself does not guarantee a grasp of musical relationships. A range of musical instruments and contexts must be explored so that young musicians become able to coordinate their skills across a range of situations.

In the absence of musical education, patterns of taste remain stable throughout the course of one’s life (Zenatti, 1993, p. 185). Indeed, too many adults consider themselves to be “unmusical.” They might be encouraged if results of a musical ability test showed they were in fact average or even superior (see also Gibbons, 1982). Older adults can benefit greatly from music participation and instruction (Darrough & Boswell, 1992). Can do is not always the same as will do. Motivational factors are undoubtedly of vital importance.

Whatever the future effects of technology (see Deutsch, 1996), music will always require a high level of cognitive ability and commitment.

REFERENCES


