

## The Legacy of Carl E. Seashore\*

Edwin E. Gordon, Distinguished Professor In Residence  
University of South Carolina

In 1919, Carl E. Seashore was the first to publish a standardized battery of music aptitude tests. It was called the *Seashore Measures of Musical Talent*. The tests were based on approximately 20 years of research into the characteristics and description of music aptitude. With assistance from students, and particularly his colleagues Joseph Saetveit and Don Lewis, Seashore's research continued, and in 1939 the battery was revised and renamed the *Seashore Measures of Musical Talents*. By adding the "s" to the word "talent," Seashore emphasized his disagreement with those who contended that there was only one overall music aptitude. In the 1919 and 1939 editions, there were two series, A and B, the only difference between the two being that the content, not the design, of B was more difficult. The Columbia Graphophone Company produced three 78 rpm records for the 1919 version, and the RCA Manufacturing Company reproduced them in a 1939 version. The test manual for the revision was published by the University of Iowa. Though the tests remained unchanged, in 1957 one 33-1/3 rpm record of Series A of the 1939 version, published at that time in conjunction with a test manual by the Psychological Corporation, replaced the 78 rpm records. Series B was discontinued without explanation. Seashore's research into the description of music aptitude culminated in the test battery.

Seashore distinguished between the psychological aspects of sound, which he called pitch, loudness, time, and timbre, and the corresponding physiological aspects of sound, which he called frequency, amplitude, duration, and form. He reasoned that the physiological aspects are the bases of the psychological aspects, and that unless one had the acuity to understand the physiological, one would be unable to comprehend the psychological. Thus, for Seashore, music aptitude was best described by what can be observed objectively in the sound wave. He went on to say, however, that the four

physiological sensory capacities and their psychological counterparts, the sense of tone quality, the sense of volume, the sense of rhythm, and the sense of consonance, are best understood when they come together and function as “complex forms.” Seashore declared that to derive a comprehensive description of music aptitude, these complex forms must be considered without singling out their elemental components, the physiological and psychological underpinnings.

Although Seashore believed in the importance of stressing the complex forms of music aptitude, his research focused on the acoustical (the physiological and psychological) aspects of music aptitude. Evidently Seashore believed that in order to understand the complex, one must first understand the simple, and he left it to others to investigate the more elaborate structure of music aptitude. Nonetheless, Seashore gave direction to the research of those in generations to follow by making the need obvious. Without his keen mind and intellectual curiosity to identify important ideas in conjunction with his fundamental research to serve as an impetus and a guide, I believe our understanding of music aptitude would not have progressed to the extent it has.

The basis of Seashore’s description of music aptitude can be understood by examining his research into the simple psychological aspects of pitch, loudness, time, timbre, and consonance. And, although Seashore does not categorize them as either elemental or complex forms, his research into tonal memory and rhythm must be given commensurate consideration, because they, along with pitch, are the most widely acknowledged components of music aptitude that have come to be associated with the Seashore legacy.

For Seashore, the sense of pitch is simply the ability to discriminate between two pitches. The more pitch aptitude one has, the finer the discriminations between pitches one

should be able to make. Specifically, Seashore believed that to demonstrate pitch aptitude, one had to be able to determine if the second of a pair of pitches is higher or lower than the first. It was important that the tones representing the pitches be pure and lacking in overtones and harmonics, and that all acoustical properties other than pitch be held constant. Thus, the tones to be compared could not be produced using an actual musical instrument. An electronic beat-frequency oscillator, which at the time had just replaced resonators and tuning forks, was used to produce the tones. The two tones in the pair were performed at approximately 500 cycles, each had a duration of 0.6 of a second, and the pitches differed from each other by only 2 to 17 cycles.

Seashore described the sense of loudness as the ability to determine if the second of a pair of tones is stronger or weaker than the first. Again, to be assured of pure tones with the degree of loudness as the only variable under scrutiny, the same apparatus for measuring pitch discrimination was used for measuring loudness discrimination. The tones were held constant at 440 cycles, and they varied in loudness by as little as 0.5 to 4.0 decibels.

To measure the sense of time, Seashore again used the beat-frequency oscillator for producing pairs of pure tones. The frequency of the tones was, as before, held constant at 440 cycles, and the differences in duration between them ranged from .05 to .30 of a second. The capacity to discriminate time was determined by the ability to know whether the second of the pair of tones was longer or shorter than the first.

Timbre discrimination was measured by the capacity to determine whether the second of two tones produced with a special generator sounded the same as or different from the first. Each tone consisted of a fundamental component with a frequency of 180

cycles and its first five overtones. The timbre of the tones was varied by reciprocal alterations, ranging from 0.7 to 10.0, in the intensities of the third and fourth harmonics.

Seashore considered the sense of consonance to be an important component of music aptitude. He experienced difficulty, however, in measuring that capacity. The directions for responding to stimuli could not be made clear, and subjects who participated in the research made preferential rather than objective judgments. Seashore reasoned that to demonstrate a sense of consonance, one had to be able to perceive which of a pair of dyads had more “smoothness, purity, and blending.” Based on Constantine F. Malmberg’s research, the order of merit of two simultaneously sounding tones on a consonance-dissonance scale that Seashore identified were, from highest to lowest, the octave, perfect fifth, major sixth, major third, perfect fourth, minor sixth, minor third, diminished fifth, minor seventh, major second, major seventh, and minor second. Interestingly, it was found that whether the tones were selected from the tempered or untempered scales and performed using a resonator, piano, or pipe organ had no affect on the judgments (preferences) reported by the subjects.

The five components of music aptitude that Seashore postulated were observable in the sound wave. That was not the case for the two capacities, tonal memory and rhythm, the latter being included in the 1939 revision of the battery. Not only did Seashore not discuss their allusive quality in the sound wave, he did not refer to either as a category of elemental or complex forms. Nonetheless, he considered the two to be as important as the other five capacities in contributing to the description and measurement of music aptitude.

Seashore proposed that to demonstrate tonal memory, one had to be able to determine by number which single pitch in a series of three, four, or five is different when the series is heard twice. Christian Paul Heinlein, however, offered evidence that suggested

this might not be the case, because he discovered that persons known to be “musical” apprehend a series of pitches as a totality and find it disturbing to have to concentrate on individual pitches. Seashore, nevertheless, disregarded such criticism and declared only that it is important that the pitches in either series have no syntactical relation to one another, because, as he stated, if they did, what would be measured would be more indicative of music achievement than of music aptitude. Thus, because of the “atonal” or neutral nature of the series, the measures could be used to undertake valid research with all persons, regardless of their race, religion, or nationality. In a word, the series of pitches would not be indigenous to any particular culture and so would be culture free. The eighteen tempered chromatic steps upward from middle “C” performed on an electric organ served as the stimuli, with tempo and intensity carefully controlled. No explanation was offered about why pitches 100 cents apart (chromatic steps) were used in the tonal memory subtest, whereas much finer differences were incorporated into the pitch discrimination subtest.

Rhythm patterns served as the stimuli in Seashore’s rhythm subtest. One had to be able to determine whether two rhythm patterns consisting of five, six, or seven durations sounded the same or different. A beat-frequency oscillator set at 500 cycles was used to perform the rhythm patterns, and the tempo was held constant at 92 quarter-notes per minute. According to the short accompanying test manual for the 1957 version of the test battery, “The first ten items contain patterns of 5 notes in 2/4 time; the next ten, patterns of 6 notes in 3/4 time; and the last ten, patterns of 7 notes in 4/4 time.” Because the patterns had no melody and no accented beats, all patterns were aurally perceived by most subjects in duple meter, regardless of the intention of the test author or the way the patterns were notated. Further, because even as Seashore suggested, humans subjectively organize for themselves series of objectively produced sounds into patterns, such as the ticking of a clock or the sound of moving train wheels, this tends to reinforce the idea that the rhythm

subtest may not be culture-free either. It is curious, meanwhile, that to demonstrate the capacity of tonal memory, one had to count the sequential position and be acutely aware of individual pitches in a tonal pattern, whereas to demonstrate the capacity of rhythm (actually rhythm memory), one had only to compare one complete pattern to another in terms of sameness and difference. Seashore never addressed why this should have been the case, although over time it has become a very important issue. Perhaps even more problematic is that although Seashore made sure it was understood that pitches had to be musically unrelated in the tonal memory subtest in order not to confuse tonal achievement with tonal aptitude, he did not make clear how it was possible that rhythm aptitude would not be confused with rhythm achievement if meter were subjectively established by the listener in the rhythm subtest.

Although the sense of consonance played a formidable role in Seashore's early deliberations about music aptitude, he later took the position that the sense of timbre was the more important factor of the two in the description of music aptitude. Moreover, although initially he was of the opinion that the sense of time was fundamental to music aptitude, he soon came to the conclusion that it was not sufficient in itself and that a rhythm component was essential for fully acquiring an understanding of music aptitude.

Because Seashore was a great pioneer, the first researcher to give sustained scholarly attention to the nature and description of music aptitude, he had many critics in the United States, Canada, and abroad. That he changed some of his positions was undoubtedly due more to the views of his adversaries than to his advocates. Christian Paul Heinlein, on the basis of his own research, published two studies in 1928 in which he criticized Seashore's philosophical and experimental methodology and conclusions. James L. Mursell, coming from mainly a philosophical orientation, disagreed with Seashore's

validation procedures and his interpretation of data. The objections in England by Harry Lowery and James Mainwaring, for example, were even greater.

It was difficult for Seashore's critics to understand how the discrimination of pitches less than a semitone apart, and particularly those that differ only nine cents of the one hundred cents in a semitone, had a significant relation to music aptitude. Though it might aid in tuning an instrument, they reasoned that pitch discrimination would seem to offer little in helping a musician learn how to adjust intonation in ensemble performance. Moreover, Seashore's antagonists did not agree that time discrimination could contribute in any significant manner to rhythmic understanding, that loudness discrimination had any relevance to expressive understanding, or that timbre discrimination could in any way be connected with the sense or production of tone quality. These theorists took the position that the capacities to discriminate between tones in terms of pitch, time, loudness, and timbre are measures more of acoustical acuity than of music aptitude. Specifically, they believed that one's ability to make generalizations about a series of musically unrelated pitches would have little if any value for predicting how well one can develop a sense of tonal syntax in music.

These concerns, as well as others, led to the Gestalt-atomistic controversy. The essence of the argument by the critics was that music aptitude cannot be divided into component parts, and that even if it could, the whole would be different from, if not greater than, the sum of its parts. Seashore maintained the position that indeed music aptitude does comprise several different capacities, and that in order to gain a precise understanding of the totality of music aptitude and to fully describe it, each of the component parts of music aptitude must be considered independently of every other and studied in that manner. For example, on the basis of the near-zero correlations among all of his subtests, Seashore insisted that pitch and tonal memory, on the one hand, and time and rhythm, on the other,

are distinctly different types of capacities from each other. To reaffirm his position, Seashore refused to develop norms for a composite score on his test battery, believing that the value of describing and measuring the component parts of music aptitude would be obscured by the acknowledgment of an all-inclusive total score.

The Gestalt-atomistic controversy went further, however. Seashore's critics believed that in order to arrive at a true description of music aptitude, tests must include musical rather than acoustical content, and that the medium through which the music is performed must be that of one or more musical instruments. Seashore protested against this view, maintaining that just as an artist must possess visual acuity, that without the acoustical capabilities that his tests measure, a musician would not adequately be able to comprehend music itself. There was no question in Seashore's mind but that the physiological forms the basis for the psychological in fusing elemental capacities into a complex musical whole. He did admit, however, that "it does not follow that goodness in these capacities alone will make a good artist" (*Seashore Measures of Musical Talents*, 7). He was quick to admit that success in music was also dependent on factors that his tests did not measure, perhaps personality traits, motor ability, and the quality of music instruction. Such an admission prompted some to observe that the Seashore tests have negative validity, and that because they can be used to indicate only what cannot be included in a description of music aptitude, they serve little purpose in identifying what should be included in that description.

Other of Seashore's concepts that came under attack had to do with the manner in which he distinguished between music aptitude and music achievement. As I have explained, only when measuring timbre and rhythm did Seashore ask whether what was to be compared was the same or different, and thus he argued that the tests required absolutely no musical achievement on the part of the subjects being tested. The critics had little quarrel



with that. What disturbed them, however, was that in the case of the pitch test, one had to know the difference between higher and lower, and that knowledge of those words required music achievement. They explained that those words are abstractions, only to be understood in terms of the placement of notes on the staff, and it could not be denied that music literacy, specifically reading, is an integral aspect of music achievement. While it can be said that familiarity with the difference between stronger and weaker and longer and shorter does not depend on music achievement, using those words in association with the sound of music might require knowledge beyond that which one might have if one had no exposure to music instruction of any kind. With regard to tonal memory, it was the British, in particular, who observed that to count and remember the ordinal number of a pitch that was changed in a series of three or more pitches, one necessarily had to have acquired some degree of academic achievement, not to mention a substantial degree of intelligence. To that extent, they claimed, Seashore was unwittingly concurring with the view that music aptitude and intelligence are very closely related.

The absence of preference measures in Seashore's thinking loomed large among the Europeans. Those who held to the Gestalt position were particularly unyielding in their view that musical preference was the *sine qua non* of music aptitude. That alone, some believed, would best distinguish among those with high, average, and low music aptitudes. Seashore found himself in a difficult position, because neither the sense of consonance nor the sense of timbre was designed to yield preferential responses. Yet, even Seashore acknowledged that his measure of the sense of consonance was problematic, because those whose music aptitude he intended to measure based their responses on their preferences. After all, is it possible to expect that being told to react to smoothness, purity, and blending would not result in a preference rather than in objectivity? Still, for all intents and purposes, Seashore was not among those who could be called upon to defend the importance of musical preference to music aptitude.

There was also a problem associated with Seashore's technique for measuring rhythm aptitude. He used very short durations, of always the same length, excluding any sustained tones. Thus, the difference between silence (rests) and duration in terms of intended note lengths had to be imposed subjectively by the listener. That listeners did just that compounded the problem of normal illusion, an acoustical phenomenon that permeated all of Seashore's measures. For example, high pitches subjectively sound louder than low pitches. Thus, even though it was thought that pitch was the only variable being measured in the pitch subtest, a listener with a poor sense of pitch and a good sense of loudness could nonetheless score high on the pitch subtest simply because of the association of the normal illusion of loudness with the higher pitch in a pair. There is also the normal illusion that a pitch is lower than it actually is if it is long or rich in overtones, and this too would affect the validity of the time and timbre subtests. The fact that in Seashore's tests normal illusion could not be avoided gave justification and impetus to the Gestalt theory of music aptitude.

All of the aforementioned criticisms aside, some researchers questioned Seashore's assumptions, not on the basis of the musical content and techniques he used, but rather on his steadfast conviction that music aptitude could not be improved with practice and training. They believed that by contradicting Seashore's conviction in this matter, they might shed light on the credibility of his description of music aptitude. Researchers on both sides, such as Edward H. Cameron, Alexander A. Capurso, Earle Connette, James Mainwaring, James Martin, Robert H. Seashore, Franklin O. Smith, and Elizabeth Taylor, conducted numerous studies to determine if Seashore's assertion would or would not be corroborated objectively. A classic study and perhaps the best known, one which Seashore's son, Robert, with the encouragement of his father, helped design, is Ruth Wyatt's. Though the experimental designs of many of the studies would not measure up to today's standards and the findings were often contradictory, most probably because such

an array of inappropriate as well as appropriate criteria were used to investigate the validity of the tests, the consensus was that scores on the tests, particularly the measure of one's sense of pitch, could indeed be improved with practice and training. Unfortunately, the researchers made no distinction between the nature of practice as opposed to the nature of training, nor was it apparent that they were aware that when different types of practice and different types of training are examined, conflicting results should be expected. Moreover, they relied on statistical significance alone, giving no attention whatsoever to practical significance, when they compared the scores of students who received practice and training to those who did not.

Nevertheless, Seashore's critics declared that his tests were measuring music achievement and that, therefore, the results were irrelevant to the description of music aptitude. In his own defense, Seashore explained that the instruction provided in the experiments had raised students' cognitive scores and not their physiological scores, and that physiological results would be obtained if students were tested individually in a laboratory, whereas cognitive results are naturally obtained when students are tested in a group in a large room with inadequate supervision. He also contended that a rise in cognitive scores is only temporary, and that if students were retested some time later, it would be discovered that the initial increases in their scores would disappear. Given present understanding of research techniques, there is good reason to believe that Seashore was correct. Unfortunately, none of the groups was ever retested.

Meanwhile, there were those who believed that the 1919 *Seashore Measures of Musical Talent* was a music achievement test because the B series was purposely designed to be more difficult than the A series. That may seem logical at first, but upon reflection it is not reasonable. Because one test is designed differently and may be more difficult than another, it does not necessarily follow that both are achievement tests, or that the easier one

is an aptitude test and the other is an achievement test. As will be explained later, different levels of complexity of musical content, as well as procedures for performing that content, are necessary to measure the developmental music aptitudes of students who have had varying amounts of exposure to music, and this does not jeopardize the validity of the tests as measures of music aptitude.

Of primary importance to music psychologists who are familiar with the concept of audiation (the ability to hear and understand music for which the sound is not, or may never have been, physically present) is that Seashore was concerned with only the first stage of audiation in the description, and ultimately the measurement, of music aptitude. Whether Seashore made that decision deliberately is open to debate. Recognition of the types and stages of audiation is recent (the word “audiation” was coined some 20 years after Seashore’s death), and it is unknown from his writings whether Seashore went beyond the general concept of imagery to give consideration to the possibility of the existence of stages, if not types, of audiation that we progress through when we learn music. Nonetheless, Seashore obviously believed that music aptitude is best described by how well one can retain and compare in aftersound the immediate aural impressions made by tones that were heard seconds before. Because the two tones used to measure the sense of pitch, loudness, time, and timbre and the series of three to five tones used to measure tonal memory are without musical syntax, audiation at the second and higher stages was not measured. As I have explained, developmental music aptitude includes the first and second stages of audiation, and stabilized music aptitude includes the first, second, and third stages. That only the first stage is considered in Seashore’s research indicates that although in his description of music aptitude he speaks of the importance of both the musical and the acoustical aspects, in his measurement of music aptitude, the results of which were to be used to support his position, he measures only the acoustical aspects.

It is true that the series of tones, which are essentially rhythm patterns, in the rhythm test are syntactical. It would be an exaggeration, however, to say that even that one subtest of the six in the revised battery describes and is a valid measure of music aptitude, because all of the patterns are audiated in the same meter. Again, even though the rhythm subtest may be the most indicative of music aptitude when compared to the other subtests, Seashore considered the sense of time, not rhythm, to be basic to music aptitude, primarily because it can be observed in the sound wave.

Regardless of how one may respond to Seashore's description of music aptitude, the general validity of his objective tests intended to support his views must be examined. Unless the validity of the tests can be demonstrated, agreement or disagreement over the accuracy of his description of music aptitude becomes a moot question. If it can be shown that the tests have validity, his description of music aptitude would appear to have validity, but if no validity can be found, it is almost certain that his description of music aptitude is inadequate.

The reliability of the tests will be considered first, because this, as well as some other factors, is an indirect, if not a preliminary, indication of the validity of the tests. Unless results on a test are consistent, they cannot be valid. Seashore reported the reliability for each of his subtests in the 1939 revision in the test manual that accompanied the 1957 edition. For 300 students approximately 10 and 11 years old in grades 5 and 6, the Spearman-Brown corrected odds-evens internal consistency reliabilities range from .69 for the timbre subtest to .84 for the tonal memory subtest. The remaining four subtest reliabilities ranged from .73 to .79. For 300 students approximately 12 and 13 years old in grades 7 and 8, the Spearman-Brown corrected odds-evens internal consistency reliabilities ranged from .69 for the rhythm subtest to .87 for the tonal memory subtest. The remaining four subtest reliabilities ranged from .75 to .80. The Spearman-Brown corrected odds-

evens internal consistency reliabilities for senior high school students, college students, and adults together go no higher than .88 for any of the subtests.

It is unfortunate that the reliability coefficients were calculated for two or more grade or age levels combined, because the resultant coefficients become spuriously high when heterogeneity is artificially increased without objective justification. The more variability in a group, the higher the reliability that can be expected. Thus, though the reliabilities for separate grades and ages are not reported by the test author, it can be reasonably assumed that they are lower than those for the combined grade and age groups.

The relationships among most of the subtests were found to be low. That finding led Seashore to maintain that each subtest was measuring one of the factorially pure music capacities he had described. It should be noted, however, that a correlation between any two measures is limited by the low reliability of even one of the measures and to a greater extent by the low reliability of both.

Seashore, disregarding the diverse composition of the groups he used in gathering norms, not only portrayed the reliabilities of his test as high, but he pronounced the tests valid as a result of those reliabilities. He took the liberty of interpreting the internal consistency reliability coefficients as “internal validity” coefficients. To Seashore, internal validity meant internal consistency. His circular reasoning was that because the two halves of each subtest were highly related, they were measuring the same capacity. In other words, no external criterion of validity was needed because each half of the subtest validated the other. Yet the problem remained: Seashore had no evidence that either half of the subtest related to a criterion representing music aptitude.

The reliabilities reported by Seashore are high enough to indicate that his subtests could demonstrate experimental validity, but the validity of a test may only be inferred, not proved, from reliability data. Nonetheless, he claimed that all of the subtests “have been validated for what they purport to measure. When we have measured the sense of pitch, that is, pitch discrimination, in the laboratory with high reliability and we know that pitch was isolated from all other factors, no scientist will question but that we have measured the sense of pitch”. Seashore summarized his position by stating that attempts to validate his subtests by correlating them with “fallible external criterion measures such as judgments of omnibus musical behaviors were inappropriate”.

I can empathize with Seashore’s frustration in identifying an external valid criterion measure. Just as the magnitude of the correlation between two tests is dependent on the reliability of both tests in question, so, too, is the magnitude of the correlation between a test and a validity criterion dependent on both of those factors. From personal experience, I have observed that students’ performances as well as teachers’ judgments of their students’ musicianship are inadequate indicators of the students’ music aptitude. Further, even a heterogeneous group of subjects is generally so lacking in variability in music achievement that it is impossible to expect them to provide useful responses. For example, because audiation cannot be directly observed, the most acceptable indirect criterion that might be used to validate a music aptitude test would be samples of students’ improvisations. Yet few persons in our society are able to improvise even at the most elementary level.

Validity, particularly of an aptitude test, is understood best through longitudinal predictive studies. Though a predictive test is not necessarily an aptitude test (although a valid aptitude test must be a predictive test), customarily the validity of an aptitude test is determined through its predictive validity. Longitudinal predictive validity studies are rare

because they are more expensive and time consuming than concurrent or criterion-related validity studies.

A longitudinal predictive study undertaken by Hazel Stanton for the Seashore battery (commonly referred to as the Eastman Study, because data were collected from students at the Eastman School of Music over a period of 10 years) was given a poor reception by reviewers, because Stanton designed her study in an unorthodox manner and then reported her results in a way that deprived readers of some of the most important findings. To predict success in music, she took the composite score of all five Seashore subtests included in the first edition (pitch, loudness, time, tonal memory, and consonance) and combined them with other criteria, including teachers' ratings, scores on an entrance audition, numerical values for past music achievement and academic achievement, and intelligence test scores. By combining those scores, the power of the Seashore tests themselves to predict the subjects' rate of graduation from the conservatory was obscured.

The following are the specific details of the design of the Stanton study. In 1925, 164 freshman entered the Eastman School of Music. Upon admission, each of the students was labeled "safe," "probable," "possible," "doubtful," or "discouraged" in terms of a prediction as to whether the student would graduate from the school in 1929, within the normal four-year period. Students were put into one of those five categories on the basis of the combined criteria described above, which Stanton referred to as the "cumulative key." Although other criteria of success were examined, such as scholarships, honors, and recital appearances, graduation from the school was given primary consideration. It was found that 60 percent of students in the "safe" group, 42 percent in the "probable" group, 33 percent in the "possible" group, 23 percent in the "doubtful" group, and 17 percent in the "discouraged" group graduated in 1929. It can only be assumed that the percentages for all of the Seashore subtests alone in predicting success in music would have been even less



accurate, although simple biserial correlations between the total scores on the Seashore battery and graduation would have been helpful in evaluating the validity of Seashore's description of music aptitude. Even if those coefficients proved to be low, however, it would have been understandable, because there are many extra-musical factors involved in determining whether or not one graduates from school, financial resources being not least among them. Nonetheless, specific information pertaining to the study has probably long since been destroyed.

Seashore gave musicians and educators a broad basis for their initial thinking about the nature and measurement, as well as the description of, music aptitude, although the realm of his contributions has generally been relegated to subjectivity and not considered in terms of the objectivity he championed. Some of his students and others he influenced indirectly duplicated his work, while others went beyond it. Among his more progressive students was Jacob Kwalwasser. In 1931, he and Peter Dykema published the *Kwalwasser-Dykema Music Tests*, which may have stimulated Seashore's interest in revising his test battery. Meanwhile, research by Herbert Wing, in England, gave support to the idea that music preference plays a formidable role in the description of music aptitude.

Much has been learned about the nature and measurement of music aptitude since Seashore's time. Most important, perhaps, is the concept of developmental music aptitude in contrast to stabilized music aptitude, and the role audiation serves in understanding the distinction between the two. It was Seashore's early insights that established the guidelines for such current research.

---

\* Freely adapted from my book *Introduction to Research and the Psychology of Music*, Chicago: GIA, 1998.

## Abstract

Carl Emil Seashore was the first American to become interested in the objective measurement of music aptitude. His early research culminated in the publication of the *Seashore Measures of Musical Talent* in 1919. Seashore's ideas about music aptitude gave rise to great controversy, nonetheless they laid the groundwork for future researchers. To truly understand the contribution of this great pioneer, his concepts must be understood in detail. That is probably best accomplished by becoming familiar with the background and content of the Seashore's test batteries.