

# What Does it Mean to be Musical? A Critical Evaluation of Recent Instruments Designed for the Measurement of Musical Ability in the General Population

Bethany Weston  
*Durham University, BA Music*

## ABSTRACT

Assessments of musical ability are often based on the measurement of musical training, musical expertise, and the ability to produce music. Often, assessments of musical competence overlook musical aptitude or innate ability, and as a result, assessments cannot provide a true representation of musical ability in the general population. This study considers some of the various definitions of ‘musical ability’ and the ways in which it is measured, before evaluating three recent instruments designed specifically for the measurement of musical ability in the general population. This paper focuses on the *Profile of Music Perception Skills (PROMS)*, the *Goldsmiths Musical Sophistication Index (Gold-MSI)*, and the *Music Use and Background Questionnaire (MUSEBAQ)*, exploring the design of each test battery and considering the effectiveness of each assessment.

## INTRODUCTION

In Western culture, there is a common understanding that some people are musical, and others are not. Generally, assessments of musical competence are based on the level of musical training people have received (Law & Zentner, 2012), and an individual’s degree of musicality is often judged by their ability to produce music, whether this be via performance on an instrument or voice, or via the composition or arrangement of music (Levitin, 2012). Often, little or no distinction is made between musical expertise (musical ability that has been acquired via musical training and experience), and musical aptitude (innate or natural musical ability), and by consequence, assessments can overlook “musically untrained individuals with musical skill, as well as those who, despite extensive musical training, may not be as skilled” (Law & Zentner, 2012). A number of studies have been conducted regarding the measurement of musical skill, and a number of instruments have been designed to assess musical ability in the general population. This review will consider the varying definitions of ‘musical ability’ and the ways in which it is measured, before evaluating recent tests that have been designed for the assessment of musical ability in the general population. This study will focus particularly on the *Profile of Music Perception Skills* or *PROMS* (Law & Zentner, 2012), the *Goldsmiths Musical Sophistication Index* or *Gold-MSI* (Müllensiefen, Gingras, Musil & Stewart, 2014), and the *Music Use and Background Questionnaire* or *MUSEBAQ* (Chin, Coutinho, Scherer & Rickard, 2018).

## 1. WHAT IS ‘MUSICAL ABILITY’?

Defining musical ability can be a challenging task due to the variety of attitudes regarding what it means to be musical. Some of the most obvious signposts for musical competence include the ability to sing or play an instrument to a relatively high standard, the ability to compose original music, the ability to demonstrate a comprehensive understanding of music theory, and the ability to read music to name only a few. All of the aforementioned factors are elements of music that can be learned, and thus, they can be categorised under musical ‘expertise’. The measurement of musical expertise certainly provides an indication of musical ability, but what about the elements of musicality that we are born with? According to Levitin (2012), “musical ability is popularly regarded to be innate: one either is or is not born with musical talent” (p. 633). This is supported by Jakubowski, Müllensiefen and Stewart (2017) via their suggestion that latent absolute pitch (the ability to recall the original pitch level of familiar songs) is widespread, and that this ability can develop in children from as young as four years old and continue into adulthood. Throughout his study entitled *What Does it Mean to be Musical?* Levitin (2012) outlines a number of ways in which musical ability extends beyond the realms of performance and composition, claiming there is a need for “an inclusive approach to capture as many musical behaviours as possible in initial studies of understanding what it means to be musical” (p. 635). Then, regarding the assessment of musical ability, he states:

we need to be more sensitive to the variety of ways that assessing musicality can present itself, such as in production and perception, and technically and emotionally. Assessments need to allow for spontaneity and creativity. (p. 635)

Levitin’s suggestion is supported by Hallam and Prince (2003) who claim that “musical ability [includes] a wide range of perceptive activities such as listening, appreciating and responding to music”, and thus, assessments must account for this (p. 17). Levitin (2012) illustrates this as he provides excellent examples of people who demonstrate unexpectedly significant levels of musical competence. One example that stands out is his discussion of DJs, during which he states:

Some DJs can listen to the briefest excerpt of a musical piece, often 1s or less, and identify the title, composer, and performers and distinguish several different performances of the same piece by the same group. DJs can introduce new connections between music we might not otherwise notice and introduce us to music we might not otherwise discover. (p. 634)

It is not uncommon for DJs to have experienced no or very little musical training, and DJs are not required to play an instrument or to compose original music. Furthermore, a successful DJ does not need to be able to read music, nor do they need to possess a particularly comprehensive understanding of music theory. Thus, if the assessment of musical ability is restricted to such constructs, as it often is in music education at least, the high level of musical competence exhibited by DJs may be overlooked. Levitin (2012) elaborates on the complex cognitive process that DJs demonstrate, stating that they must be able to listen to music whilst simultaneously “searching a vast mental repertoire of music to find a template match for chords, melodies, rhythms, timbres, or other component features”, and they must also be able to “[perform] mental transpositions to place [music] into equivalent keys and tempi” (p. 634). Levitin rightfully states that “recognising these sorts of musical connections is not something that all musicians and not even all great musicians can do” (p. 634), suggesting that DJs demonstrate a significant level of musical ability, despite often lacking formal music training and the ability to play musical instruments. In addition to DJs, Levitin references music critics, recording engineers, film music supervisors, record company talent scouts, and choreographers, claiming that such “professions require various sorts of receptive (perceptual) music skills” (p. 634), but a production-biased assessment of musical ability does not account for this. Ultimately, Levitin’s *What Does it Mean to be Musical?* (2012) highlights the importance of inclusivity in assessments of musical ability, given that “musicality presents itself in a number of ways that may be uncorrelated with each other” (p. 634).

For her paper ‘*Sort of in Your Blood: Inherent Musicality Survives Cultural Judgement*’, Ruddock (2012) conducted an experiment in which she simply had conversations with twenty self-perceived non-musicians about their reasons for feeling that they are not part of the human music community. Ruddock’s method is particularly interesting in that once she had transcribed her conversations, she analysed the participants’ statements and compiled them to compose a “ballad”, demonstrating that people’s words and phrases alone can be, in a sense, musical. Whilst the subject of the ballad’s lyrics discusses the participants’ lack of musical ability, the structure of their words and phrases suggests otherwise according to Ruddock (2012), who considers the natural rhythms and inflections of the participants’ speech an indication of musicality. Ruddock’s study provides evidence in support of innate musicality, and thus, her findings support Levitin’s (2012) argument that assessments of musical ability should be “inclusive” and “sensitive to the variety of ways that... musicality can present itself” (p. 635). Ruddock’s (2012) paper highlights the distinction between musical aptitude and musical expertise, and she goes on to draw attention to the problems that arise when aptitude and expertise are not differentiated; Ruddock references a comment made by one of the participants in her study, claiming:

people grow up with ideas that are implanted at an early age about the fact that they are not talented in this or that direction and if that’s implanted as firmly as it often has been in the past, it’s very hard to change later in life. (p. 214)

This statement emphasises a fundamental flaw in restricted assessments of musical ability; particularly in school, children are often declared to be ‘musical’ or ‘non-musical’ by their teachers, and more often than not this is based only on whether or not they can sing or play an instrument. Such labels tend to impact people’s confidence regarding music in later life, and this could result in missed opportunities as ‘unmusical’ people avoid engaging in musical activities (Ruddock, 2012). The following statement by a participant in Ruddock’s study summarises the potential impact of education on people’s musical confidence:

Perceptive teachers can add immeasurably to a child’s capacity to appreciate and/or practise [music]. Conversely: imperceptive, insensitive teachers can do immeasurable damage to a child’s interest in music and capacity to enjoy (let alone perform) it. I first realised that I couldn’t sing when I was never picked for singing at school. (p. 215)

Ultimately, assessments of musical ability that focus primarily on performance and music production can overlook other important areas of musicality, and basing one’s perception of whether or not someone is musical on such restricted assessments can result in a distorted reflection of musical competence. Being told you lack musical skill can have a significant impact on confidence, and this can result in missed opportunities as people choose not to pursue music for fear of being ‘bad at it’. It is important to consider the difference between musical expertise and musical aptitude, and natural ability should be considered before declaring someone to be ‘musical’ or ‘unmusical’. Exam boards like ABRSM have constructed standardised tests to assess areas such as music performance, and whilst such examinations are reliant on the subjective judgement of examiners, they can provide a somewhat fair assessment of one’s musical expertise. The development of comprehensive musical aptitude tests can be beneficial for the assessment of musical ability amongst the general population, and more inclusive and varied tests provide opportunities for people to discover musical abilities that may otherwise be overlooked. Instruments for the assessment of musical aptitude have educational benefits, in that they enable teachers to identify particular areas of strength and weakness amongst students and thus, they can work to help each individual student reach their full potential. Furthermore, musical aptitude tests can contribute to our understanding of “the role of musical ability in non-musical faculties, ranging from motor skills and general intelligence to language processing and socio-economic competencies, such as empathy” (Law & Zentner, 2012, p. 1). A fractioning of musical ability allows us to explore “which aspects of music correlate specifically with which other cognitive abilities” (Levitin, 2012, p. 637), and an increased understanding of this will be beneficial for answering questions such as “does music make you smarter?” (Levitin, 2012, p. 637).

## 2. PRELIMINARY STUDIES SURROUNDING THE ASSESSMENT OF MUSICAL ABILITY

*The Measurement of Musical Talent: Seashore (1915).*

In the early twentieth century, Carl E. Seashore developed the first instrument for the “objective measurement of musical aptitudes, such as pitch and rhythm discrimination” (Gingras, 2016). Seashore (1915) expressed a belief similar to that

conveyed by Ruddock when he stated: “musical talent, like all other talent, is a gift of nature – inherited, not acquired; in so far as a musician has natural ability in music, he has been born with it” (p. 129). His *The Measurement of Musical Talent* suggests that people possess varying degrees of innate musicality, and Seashore (1915) attempted to design a test that could measure this objectively, as illustrated by the following statement:

Musical talent is not one thing. To amount to anything, there must be a hierarchy of talents, sufficiently related to work together. Hierarchies of talents may present entirely different organisations in different individuals. The analysis of musical talent aims first to locate the dominant traits and then to determine qualitatively and quantitatively the composition or characteristics of each group, or hierarchy of traits. (p. 130)

One element of Seashore’s (1915) *Measurement of Musical Talent* that should be particularly credited is his attempt to design a musical aptitude test that was “equally feasible for young and old, for musical and for unmusical” (p. 133). In doing so, he designed an assessment of musical ability that could be used by a variety of people, and the inclusivity of his test held potential benefits for the study of musical aptitude amongst the general population. Furthermore, Seashore’s method should be given credit because he took into account the importance of isolating specific elements of music in order to assess specific skills; when discussing his stimuli, Seashore stated:

in this test we want to vary and measure pitch. Time, timbre, intensity, sequence-complexes and all other factors of tone must therefore be kept relatively constant or uniform; and all conditions must be kept as simple as possible. (p. 133)

For the measurement of pitch, Seashore (1915) utilised tuning forks and a resonator. Participants were played two successive tones and asked to identify which of the two was higher, starting with the largest interval of 30 vibrations and working their way through remaining intervals “in the order of decreasing magnitude” (pp. 133-134). The justification for this method is particularly interesting; Seashore discusses the phenomenon of pitch illusion, stating:

It would be no test at all merely to ask the observer if he heard a difference [between two successive pitches], as the early experimenters did; he must also be required to give the direction. By virtue of the illusions we often tend to hear two tones of the same pitch as different and sometimes feel a higher degree of certainty in a judgement which is wrong. (p. 137)

Seashore’s method was relatively straightforward, but his consideration of pitch illusion and the subsequent test design demonstrates progress regarding the measurement of pitch discrimination. In this sense, Seashore’s pioneering study was effective, and this element of his work arguably contributed to the design of later pitch discrimination tests.

Seashore (1915) acknowledged and discussed individual differences between participants and considered the impact of cognitive limitations on test performance, such as misunderstanding and lack of attention. He also considered

physiological limitations. Seashore’s *Measurement of Musical Talent* discusses important factors that may influence performance on the pitch discrimination test, including age, gender, the advantage of possessing absolute pitch, and the disadvantage of being ‘tone-deaf’. Regarding the analysis of data, Seashore (1915) reduced all numerical values to a common denominator via the method of “percental rank” (p. 143), enabling a comparison of values. This was particularly beneficial for the creation of visual representations of musical talent; Seashore created charts for individual participants, illustrating their test results for areas such as pitch discrimination, tonal imagery, tonal memory, free rhythm and regulated rhythm. The charts also incorporate information from a systematic questionnaire including vocal training, instrumental training, and musical appreciation, and this information combined with the test results conveys a clear insight into participants’ musical ability, regarding pitch and rhythm at least.

Seashore’s pioneering study laid the foundation for the development of later musical aptitude tests and a number of more recent instruments have utilised similar methods, particularly regarding the measurement of pitch discrimination. As a result of advancements in technology, recent studies have been able to create musical aptitude tests and questionnaires that can be released and completed online, and by consequence, instruments such as *Gold-MSI* and *MUSEBAQ* have the potential to reach a considerably larger sample than Seashore’s assessment. Furthermore, advancements in music technology have enabled recent studies to produce electronic stimuli that is more versatile and perhaps more accurate than tuning forks. Electronic stimuli for a pitch discrimination test, for example, can be produced in such a way that it eliminates all harmonics, enabling a pure tone to be presented at the exact pitch, volume and length required. Seashore’s *The Measurement of Musical Talent* played a significant role in the emergence and development of comprehensive musical aptitude tests, and more recent studies have expanded on Seashore’s work to produce instruments that focus on areas other than only pitch and rhythm. Prior to the production of *PROMS* (2012), *Gold-MSI* (2014) and *MUSEBAQ* (2018), a variety of tests had been constructed including numerous contributions by Seashore, the *Kwalwasser-Dykema Music Test* (1930), Raleigh Drake’s *Drake Music Aptitude Tests* (1954), Herbert Wing’s *Standardised Tests of Musical Intelligence* (1958), Arnold Bentley’s *Measures of Musical Ability* (1966), various contributions by Edwin Gordon, the *Montreal Battery Evaluation of Amusia* or *MBA* (2003), the *Clinical Assessment of Music Perception* or *CAMP* (2008), the *Musical Ear Test* or *MET* (2010), and many more (Grashel, 2008; Law & Zentner, 2012).

### 3. RECENT STUDIES SURROUNDING THE ASSESSMENT OF MUSICAL ABILITY

*The Profile of Music Perception Skills (PROMS): Law and Zentner (2012).*

Law and Zentner’s *Profile of Music Perception Skills* is a test battery designed to objectively assess musical competence amongst the general population via the measurement of “perceptual musical skills across multiple domains: tonal

(melody, pitch), qualitative (timbre, tuning), temporal (rhythm, rhythm-to-melody, accent, tempo), and dynamic (loudness)” (Law & Zentner, 2012, p.1). According to Law and Zentner:

aspects of music perception and production have been extensively investigated, [but] there has been little interest in the development of a psychometrically sound and construct-validated test capable of diagnosing individual differences in musical ability. The goal of the current research is to fill this gap. (p. 1)

In support of the suggestion that everyone possesses a degree of innate musicality, Law and Zentner (2012) suggest that “among the musically untrained, some people might reach a high level of musical proficiency if given the time and opportunity to do so”, and such people are referred to as “musical sleepers” (p.1). On the contrary, people who have received a substantial level of musical training but whose “musical proficiency languishes” are referred to as “sleeping musicians” (p. 1). The “musical sleepers” metaphor emphasises the distinction between musical aptitude and musical expertise as it suggests that musical training and experience is not the only contributing factor regarding musical competence. Furthermore, the notion of “sleeping musicians” suggests that despite extensive musical training, some individuals do not achieve the level of musical skill that is expected of them, and this implies that innate musicality plays a significant role in the development of musical ability. The *PROMS* attempts to address such issues which are often overlooked in assessments of musical ability.

The criteria for the *PROMS* test battery was as follows:

- 1) The test should be equally suitable for listeners who differ in the extent and in the type of their musical background; 2) the test should be more inclusive than previous batteries with respect to the musical perceptual components tested; 3) the test should assess each perceptual component with the greatest possible specificity; and 4) the test should meet contemporary standards for test construction in terms of validity and reliability. (Law & Zentner, 2012, p. 3)

Regarding stimuli, the materials used were somewhat elementary in nature; for the melody task, all melodies “were monophonic and composed of constant rhythms (eighth notes)”, and all were composed using a harpsichord timbre “because it is relatively neutral” (Law & Zentner, 2012, p. 4). This approach has its advantages; basic melodies or “proto-stimuli” are more “stylistically neutral” than homophonic melodies that incorporate rhythmic and timbral diversity, and thus, there is less risk of the melodies being associated with particular musical styles that may give certain participants an advantage. Furthermore, elementary stimuli “can be configured so as to test one specific perceptual skill at a time, leaving others aside” (Law & Zentner, 2012, p. 4), and this enables participants to gain a more accurate assessment of specific musical skills. However, the primary disadvantage of using elementary stimuli is that it is not always reflective of

real-world music; in everyday life, we tend to engage more with homophonic, rhythmically and instrumentally diverse music as opposed to basic monophonic melodies. Law and Zentner (2012) address this suggestion in their study via reference to the *Raven Progressive Matrices* test:

[this test] consists of abstract visual patterns that do not stand out as obvious items for the measurement of general mental ability. Thus, it is not unreasonable to expect that musical stimuli of similar parsimony would be predictive of real-life musical proficiency. (p. 4)

One element of the *PROMS* stimuli that should be credited is Law and Zentner’s consideration of potential associations and advantages; regarding pitch, Law and Zentner (2012) acknowledge that classically trained musicians may be particularly familiar with “concert pitch” (p. 6), and this may provide a slight advantage. The consideration of such factors is an important part of providing fair and reflective assessments of musical ability.

Regarding method, there are some potential problems with Law and Zentner’s approach. The tests featured in the *PROMS* utilise a same-different method; participants were presented with the initial stimuli twice, before hearing the comparison stimuli. They were then asked to select an answer from a variety of options, including “definitely same”, “probably same”, “I don’t know”, “probably different”, and “definitely different” (p. 6). The main problem with this method is one highlighted by Seashore (1915) in his *The Measurement of Musical Talent*; particularly regarding pitch tests, the phenomenon of pitch illusion may impact participants’ answers. Seashore stated “it would be no test at all merely to ask the observer if he heard a difference” (p. 137), but the same-different method utilised by Law and Zentner does just that. Perhaps this method is somewhat flawed in pitch discrimination tests as suggested by Seashore, but it may be sufficient for tests of other perceptual skills; a same-different approach is arguably well suited to comparisons of rhythm, tempo, and timbre. Furthermore, this method is relatively straightforward, and by consequence, the tests are appropriate for people of varying ages and abilities. This is, after all, one of the primary aims of the *PROMS* test battery.

Overall, Law and Zentner’s *PROMS* instrument provides a comprehensive assessment of musical aptitude that can be utilised by the general population. This is particularly possible because Law and Zentner administered a condensed version of the test that can be completed in under an hour. Whilst this version does not feature all subsets, the subsets featured “do not require the absolute silence necessary to discern subtle variations in timbre, pitch, or loudness, thereby making them more suitable to be administered online” (Law & Zentner, 2012, p. 11). Whilst Law and Zentner’s study utilised a limited sample (78 participants for study 1), future research can be conducted with relative ease due to the accessibility of the shortened test battery. One flaw of the *PROMS* test is that it does not measure the perception of expressive musical qualities, such as phrasing and balance, nor does it assess production tasks. This is, however, acknowledged by the researchers.

*The Goldsmiths Musical Sophistication Index (Gold-MSI): Müllensiefen, Gingras, Musil, and Stewart (2014).*

The *Gold-MSI* is an online test battery designed to assess self-reported musical skills and behaviours in the general population (Müllensiefen et al., 2014). Unlike the *PROMS* test, which focuses on the concept of musical aptitude, the *Gold-MSI* concentrates specifically on musical ‘sophistication’, defining this as “a psychometric construct that can refer to musical skills, expertise, achievements, and related behaviours across a range of facets that are measured on different subscales” (p. 2). Müllensiefen et al. (2014) suggest that:

high levels of musical sophistication are generally characterised by a) higher frequencies of exerting musical skills or behaviours, b) greater ease, accuracy or effect of musical behaviour patterns when executed, and c) a greater and more varied repertoire of musical behaviour patterns. (p. 2)

With this in mind, it can be understood that musical sophistication encompasses elements of both aptitude and expertise, and by consequence, the *Gold-MSI* test battery measures a wider variety of musical skills than assessments like the *PROMS* test which focus predominantly on musical aptitude. The *Gold-MSI* is designed to focus on musical ability as it occurs in Western society, and the instrument aims to measure ability in the general population, as opposed to “smaller pathological groups” such as ‘amusics’ and professional musicians (Müllensiefen et al., 2014, p. 1). The test is accessible online and thus, it is possible to acquire data from a sufficiently large sample; the initial study conducted by Müllensiefen et al. (2014) involved 147,633 participants (p. 4). Whilst the sample was sufficient in size, the mean age was 35.2 suggesting that the test’s online nature may attract a predominantly younger sample. This is a potential limitation of all online tests, and paper editions should be considered to engage maximum participation. Another risk associated with online testing is that researchers are very reliant on participants completing the test truthfully, given that conditions cannot be controlled; before completing the *Gold-MSI*, participants were required to obtain an online-identifier from the BBC, and Müllensiefen et al. (2014) acknowledged that “it was technically possible for an individual to create a second BBC-ID and to re-take the entire test” (p. 4). An element of trust is required in this sense, but the researchers also “included a question to identify a small number of re-takers” (p. 4) in the hopes of maximising data reliability. The initial procedure for the *Gold-MSI* involved participants completing a 70-item self-report inventory online, along with a brief demographic questionnaire and four musical ability tests. One particularly positive aspect of the *Gold-MSI* test battery is that the entire testing procedure can be completed in only 25 minutes, and in combination with the ease of online access, the short time limit renders the test highly suitable for completion by members of the general population.

According to Müllensiefen et al. (2014), one of the primary goals of the *Gold-MSI* study was to “develop a self-assessment instrument that can measure expertise with regard to a variety of musical activities, not only instrumental expertise” (p. 2). Whilst self-reports are advantageous in providing an insight into participants’ personal beliefs and

experiences, there are some issues that must be considered regarding this method of assessment. One fundamental issue with self-reports and self-assessment is that personal reflection can be highly influenced by self-confidence (or a lack of this), socially-influenced biases, and various personality traits. Müllensiefen et al. (2014) acknowledge this, and in order to collect objective data that is sufficiently reflective of participants’ musical ability, they compared “scores from the *Gold-MSI* self-report inventory with [participants’] performance on Gordon’s *Advanced Measures of Audiation (AMMA)*” (p. 8). The combination of self-assessments, demographic questionnaires and musical ability tests render the *Gold-MSI* instrument a comprehensive and sufficiently objective assessment of musical ability.

Regarding stimuli, the *Gold-MSI* incorporates “tasks for assessing musical skills that are more akin to real-world skilled listening behaviours” (Müllensiefen et al., 2014, p. 2), unlike the *PROMS* test which utilises artificially created stimuli; for the beat perception task, participants had to listen to eighteen short instrumental excerpts that were “overlaid with a metronome-like beep track” (Müllensiefen et al., 2014, p. 12). One element of the *Gold-MSI* test that bears resemblance to the *PROMS* battery is the method utilised for the assessment of melodic memory; Müllensiefen et al. (2014) advocate a same-different comparison of melodies, and much like Law and Zentner’s study (2012), the melodic memory test featured in the *Gold-MSI* study incorporates “confidence ratings” to assess the participants’ level of understanding (p. 11-12). The simplicity of this method renders it easy to understand, making it suitable for use by people of varying ages and abilities. The *Gold-MSI* study does go one step further than the *PROMS* test as it also includes research surrounding socio-demographic variables. Overall, the test provides an approach to the assessment of musical ability that is suitable for use by the general population.

*The Music Use and Background Questionnaire (MUSEBAQ): Chin, Coutinho, Scherer, and Rickard (2018).*

The *MUSEBAQ* resembles the *Gold-MSI* in that both incorporate a self-report approach to the assessment of musical ability. However, whilst the *Gold-MSI* maintains focus on the concept of musical sophistication, the *MUSEBAQ* is designed to focus on music engagement; Chin et al. (2018) describe the *MUSEBAQ* as a “multi-modular self-report instrument... [designed] to measure a diverse set of music engagement constructs”, including “music background and capacity, music preferences, and motivations for using music” (p. 376). The *MUSEBAQ* “combines elements of several previous music questionnaires but combines them into a single modular instrument” (p. 376), and in doing so, the questionnaire offers an accessible and relatively comprehensive method of assessing musical engagement. Much like the *Gold-MSI*, the *MUSEBAQ* can be accessed online making widespread distribution possible. However, as is the case with all online tests and questionnaires, conditions cannot be controlled so researchers are reliant on participants’ honesty. Furthermore, the online distribution of the *MUSEBAQ* resulted in a predominantly younger age demographic, and it should be noted that this was acknowledged by the researchers; Chin et al. (2018) stated:

the age bias in each of the three studies reported means that findings must be interpreted with some caution... Replication of this work is nonetheless recommended, although given the demonstrated difficulty in attracting older participants to be involved in online research, paper and pen administration may be worthwhile. (p. 397)

One significant positive of the *MUSEBAQ* is that the finished questionnaire can be completed in less than ten minutes, and as a result, the questionnaire provides a convenient way of assessing musical engagement in the general population. Furthermore, Chin et al. (2018) claimed that the *MUSEBAQ* “is a comprehensive, modular instrument that can be used in whole, or by module as required” (p. 393). The modules include “Musicianship, Musical Capacity, Music Preferences, and Music Use Motivations” (p. 379), and “the individual modules can be used in isolation if a more targeted assessment is required” (p. 393).

Whilst the *MUSEBAQ* has been said to provide “substantially more information about an individual’s musical engagement than has been previously available” (Chin et al., 2018, p. 393), it should be noted that the questionnaire is perhaps most helpful when used in combination with other assessments of musical ability; Chin et al. (2018) stated “researchers investigating musical processing, abilities or behaviours can use the *MUSEBAQ* to complement... other chosen instrument(s)” as this enables a more thorough investigation of musical ability (p. 379). Overall, the *MUSEBAQ* provides an accessible and comprehensive tool for the measurement of musical engagement, and consequently, this instrument contributes to ongoing research surrounding the most appropriate way to measure musical ability in the general population. Chin et al. (2018) stated:

it is crucial to obtain a broader picture of the ways in which individuals use music, and how a constellation of factors, incorporating functions, processes, motivations of music engagement, sensitivity and personal commitment towards music, as well as preferences of music genre, needs to be measured and considered together. (p. 379)

The *MUSEBAQ* attempts to do just that in order to “provide a comprehensive yet concise music engagement profiling tool for individuals” (Chin et al., 2018, p. 379).

#### 4. CONCLUSIONS

Since the construction of Seashore’s pioneering musical aptitude test in 1915, considerable progress has been made regarding the development of tests to assess musical ability in the general population. Tests such as the *PROMS* (2012), the *Gold-MSI* (2014), and the *MUSEBAQ* (2018) consider elements of musical ability beyond the realms of music production and expertise, and emphasis on musical aptitude has contributed to an increased awareness of innate musicality amongst both the ‘musical’ and the ‘non-musical’. The three tests discussed in this paper provide the opportunity for accessible and comprehensive assessments of specific elements of musical ability; the *PROMS* focuses on musical aptitude, the *Gold-MSI* focuses on musical sophistication, and the *MUSEBAQ* focuses on musical engagement. Arguably, the *Gold-MSI* offers the most inclusive assessment, given that musical sophistication encompasses elements of musical

aptitude, expertise, and musical engagement (Müllensiefen et al., 2014). However, it should be noted that none of the aforementioned test batteries assess music production; none of them require participants to sing or play an instrument, for example. Consequently, the assessment of musical expertise is heavily reliant on self-assessment and information submitted via questionnaires. Furthermore, little emphasis is placed on participants’ reactions to the expressive and emotive elements of music, particularly in the *PROMS*, which does not consider how participants interpret phrasing and other expressive features (Law & Zentner, 2012). As a result of such limitations, it is perhaps best to use a combination of test batteries when assessing musical ability in the general population, at least until a gold-standard test has been established; various tests focus on different elements of musicality, and thus, different instruments can complement one another. The completion of numerous assessments would provide a more comprehensive and reflective overview of musical ability than an individual test.

Research surrounding the measurement of musical ability holds numerous educational benefits. To limit discrimination against people with no or little musical training, people should be assessed on areas other than just musical expertise before they are branded as ‘musical’ or ‘unmusical’. Assigning someone the ‘unmusical’ label can be detrimental to their confidence, and judging people purely on their level of musical training can result in missed opportunities and undiscovered potential. Assessments of musical ability should be “inclusive” to account for the various ways that musicality presents itself (Levitin, 2012). Inclusive and comprehensive assessments enable teachers to identify specific strengths and weaknesses, and by consequence, teachers can target individual skills to help students maximise their potential. In addition to this, comprehensive assessments of musical ability enable researchers to “understand the role of musical ability in non-musical faculties, ranging from motor skills and general intelligence to language processing and socio-emotional competencies, such as empathy” (Law & Zentner, 2012, p. 1). An increased understanding of the relationship between musical and non-musical faculties “might... be relevant to the understanding of deficits in [various] domains” (Law & Zentner, 2012, p. 1), and this may provide opportunities for the creation of treatment and remedies. Further research should strive to develop an objective assessment of musical ability that encompasses as many areas of musicality as possible, in the hopes of providing a reliable and truly reflective insight into musical ability amongst the general population.

#### REFERENCES

- Arnold, B. (1966). *Musical Ability in Children and its Measurement*. London: Harrap.
- Bigand, E., & Poulin-Charronnat, B. (2006). “Are we ‘experienced listeners’? A review of the musical capacities that do not depend on formal music training”, *Cognition*, 100, 100-130. doi: 10.1016/j.cognition.2005.11.007.
- Blacking, John. (1973). *How musical is man?* Washington: University of Washington Press.

- Chin, T-C., & Rickard, N. S. (2012). "The Music USE (MUSE) questionnaire: An instrument to measure engagement in music", *Music Perception*, 29, 429-446. doi: 10.1525/mp.2012.29.4.429.
- Chin, T-C., Coutinho, E., Scherer, K. R., & Rickard, N. S. (2018). "MUSEBAQ: A modular tool for music research to assess musicianship, musical capacity, music preferences, and motivations for music use", *Music Perception*, 35, 376-399. doi: 10.1525/mp.2018.35.3.376.
- Gingras, B. (2016, October). "Musicality", *serious science*, <http://seriousscience.org/musicality-7205> (accessed May 2018).
- Grashel, J. (2008). "The measurement of musical aptitude in 20<sup>th</sup> Century United States: A brief history", *Bulletin of the Council for Research in Music Education*, 176, 45-49.
- Hallam, S., & Prince, V. (2003). "Conceptions of musical ability", *Research Studies in Music Education*, 20, 2-22. doi: 10.1177/1321103X030200010101.
- Hallam, S. (2010). "21<sup>st</sup> Century conceptions of musical ability", *Psychology of Music*, 38, 308-330. doi: 10.1177/0305735609351922.
- Humphreys, J. T. (1998). "Musical aptitude testing: From James McKeen Cattell to Carl Emil Seashore", *Research Studies in Music Education*, 10, 42-53. doi: 10.1177/1321103X9801000104.
- Jakubowski, K., Müllensiefen, D., & Stewart, L. (2017). "A developmental study of latent absolute pitch memory", *The Quarterly Journal of Experimental Psychology*, 70, 434-443. doi: 10.1080/17470218.2015.1131726.
- Law, L. N. C., & Zentner, M. (2012). "Assessing musical abilities objectively: construction and validation of the profile of music perception skills", *PLoS ONE*, 7, 1-15. doi: 10.1371/journal.pone.0052508.
- Levitin, D. J. (2012). "What does it mean to be musical?", *Neuron*, 73, 633-637. doi: 10.1016/j.neuron.2012.01.017.
- Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R., & Ullén, F. (2014). "Practice does not make perfect: No causal effect of music practice on music ability", *Psychological Science*, 25, 1795-1803. doi: 10.1177/0956797614541990.
- Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). "The musicality of non musicians: An index for assessing musical sophistication in the general population", *PLoS ONE*, 9, 1-23. doi: 10.1371/journal.pone.0089642.
- Murphy, C. (1999). "How Far do tests of musical ability shed light on the nature of musical intelligence?", *British Journal of Music Education*, 16, 39-50.
- Ruddock, E. (2012). "'Sort of in your blood': Inherent musicality survives cultural judgement", *Research Studies in Music Education*, 34, 207-221. doi: 10.1177/1321103X12461747.
- Ruddock, E. (2018). "On being musical: Education towards inclusion", *Educational Philosophy and Theory*, 50, 489-498. doi: 10.1080/00131857.2016.1198248.
- Seashore, C. E. (1915). "The measurement of musical talent", *The Musical Quarterly*, 1, 129-148.
- Tan, S-L., Pfordresher, P., & Harré, R. (2010). *Psychology of music: From sound to significance*, New York, NY: Psychology Press.
- Treichler, F. R. (2013). "Testing Musical Ability: An American Dissenter and Some Related Historical Comparisons", *History of the Human Sciences*, 26, 48-68. doi: 10.1177/0952695113508121.