Personality and Individual Differences 63 (2014) 87–93

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

# Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test

Fredrik Ullén [a,*], Miriam A. Mosing [a], Linus Holm [b], Helene Eriksson [a], Guy Madison [b]

[a] Dept. of Neuroscience, Karolinska Institutet, SE-171 77 Stockholm, Sweden
[b] Dept. of Psychology, Umeå University, SE-901 87 Umeå, Sweden

ARTICLE INFO

*Article history:*
Received 21 November 2013
Received in revised form 22 January 2014
Accepted 25 January 2014
Available online 20 February 2014

*Keywords:*
Music
Sensory discrimination
Expertise
Psychological testing
Training

ABSTRACT

We examine, in 6881 twin individuals, the psychometric properties of a new test (the Swedish Musical Discrimination Test, SMDT) that was developed to tap auditory discrimination of musical stimuli. The SMDT consists of three subtests measuring discrimination of melodies, rhythms, and single pitches, respectively. Mean test taking times for the subtests were 3.0–4.6 min. Reliability and internal consistency were good with Cronbach's alpha values and Spearman–Brown split-half reliabilities between .79 and .89. Subtests correlated positively ($r$ values .27–.41). Criterion validity was demonstrated in three ways: individuals that had played a musical instrument scored higher than individuals that had not (Cohen's $d$ .38–.63); individuals that had taken music lessons scored higher than individuals that had not (Cohen's $d$ .35–.60); finally, total hours of musical training and SMDT scores correlated ($r$ values .14–.28) among those participants that had played an instrument. Lastly, twin modelling revealed moderate heritability estimates for the three sub-scales. We conclude that the SMDT has good psychometric characteristics, short test taking time, and may serve as a useful complement to existing tests of musical ability.



## 1. Introduction

Music is a human universal of profound significance for most people. At the same time, musicality, broadly defined as the capacity to learn and perform music-related tasks, seems to vary substantially between individuals. There have been many endeavours to objectively measure musicality since the early 20th century. Several standardised, explicit forms of musicality tests have been constructed, both with practical aims, such as selection of students for musical training, and for research purposes (Boyle & Radocy, 1987; Shuter-Dyson, 1999; Shuter-Dyson & Gabriel, 1981; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). Exactly how musicality is measured can make substantive differences to the information one obtains and hence to the aspects of the phenomenon one can study. Correlations between different tests and between tests and criteria such as teacher's ratings and music school grades tend to be in the range of .4–.6 (Shuter-Dyson & Gabriel, 1981). One important reason for these relatively moderate correlations is that different tests use different operationalisations of musicality. Indeed, within musicality testing there are two strong traditions, which differ in various characteristics. These are the 'atomistic' tradition of Seashore and the 'omnibus' approach of Wing (Jacobs, 1960; Seashore, 1919, 1938, 1947; Shuter-Dyson & Gabriel, 1981).

The atomistic approach is based on the assumption that musicality is made up of several relatively narrow and distinct musical abilities. This leads to an expectation of statistical independence (Gordon, 1969; Seashore, 1919) or at least low intercorrelations (Seashore, 1947) between tasks that tap into different abilities. Tests in this tradition have typically focused on basic sensory abilities, such as discrimination of various musically relevant sound stimuli. Empirical data consistently show moderate positive intercorrelations between discrimination tests (Carroll, 1993). While this to some degree supports the idea of independence of musically relevant perceptual abilities, individual differences in discrimination tasks are thus also influenced by more general factors. In fact, auditory discrimination tasks positively correlate with a broad range of non-musical cognitive tasks and psychometric modelling shows that general intelligence ($g$) is an important factor underlying the positive covariation between different 'atomistic' tests of musical discrimination (Helmbold, Troche, & Rammsayer, 2007; Lynn & Gault, 1986; Spearman, 1904; Troche & Rammsayer, 2009).

In contrast, in the omnibus approach to musicality testing, musicality is considered a general high-level ability. Tests

* Corresponding author. Address: Dept. of Neuroscience, Retzius v 8, Karolinska Institutet, SE-171 77 Stockholm, Sweden. Tel.: +46 8 52483268.
*E-mail address:* Fredrik.Ullen@ki.se (F. Ullén).

http://dx.doi.org/10.1016/j.paid.2014.01.057
0191-8869/

developed within this tradition are less concerned with characterising components of musicality but rather tend to use a holistic approach where complex, acquired musical knowledge is assessed. Typical test items may involve quality judgments of musical performances or the production of musically meaningful responses to stimuli (for example the tests of Wing and Révész; see (Jacobs, 1960)). These general differences between the two traditions also mean that the omnibus tests typically are aimed at practicing musicians, while the atomist tests can be used for a wider range of purposes. It should be pointed out that there are musicality tests that do not easily fit into either of these main traditions as they focus on musical engagement, motivation and interests rather than the cognitive capacity to process musical information as such, e.g. the Music Use Questionnaire (Chin & Rickard, 2012).

Here, a new test of musicality (Swedish Musical Discrimination Test, SMDT) is presented and its psychometric properties are analysed. The purpose of this new test is to provide measures of basic aspects of musical ability operationalised as discrimination ability for auditory musical stimuli, and the test thus continues the 'atomistic' tradition of Seashore sketched above. The SMDT consists of three subtests, Melody, Rhythm, and Pitch, which measure discrimination of rhythms, melodies, and single pitches, respectively. The three subtests are somewhat similar to the Tunes, Rhythm, and Pitch tests of Bentley (Lynn, Wilson, & Gault, 1989), the corresponding subscales of the Musical Ear Test (Wallentin et al., 2010), and the Profile of Musical Perception Skills (Law & Zentner, 2012). We aimed to design an instrument that has short test-taking time, allows for online administration, and has a suitable difficulty level for general musically untrained populations in industrialized countries. The present paper reports on the basic psychometric properties of the SMDT, including selection bias, correlations between subtests, and reliability, as well as genetic and environmental influences on each of the sub-tests based on data derived from an online administration to a larger cohort of twins. To account for the relatedness of the twins we used a randomized two-sample design, where the original sample was randomly split into two independent subsamples, in such a way that twins in the same pair were always allocated to different subsamples.

## 2. Methods

### 2.1. Participants

The participants were twins recruited from the Swedish Twin Registry (Magnusson et al., 2013). Zygosity was determined based on questions about intra-pair similarities. These have subsequently been confirmed in 27% of the twins in the registry using genotyping, showing that the questionnaire based zygosity determination was correct for more than 98% of twin pairs (Lichtenstein et al., 2002, 2006). They took the SMDT as part of a larger survey that was administered online and included numerous other questionnaires, e.g., on musical experience, personality, motivation, and interests. In total, 32,005 individuals were invited to participate, and 11,543 logged in on the questionnaire website. The present analyses are based on data from 6881 participants, i.e., the 6718 participants that completed all three SMDT subtests, as well as another 163 participants that completed only one or two of the subtests. The sample contained 1362 full twin pairs. Of these, 711 were monozygotic (MZ; identical) and 651 dizygotic (DZ; non-identical). The participants were aged between 27 and 54 (mean = 40.7, SD = 7.7); 57.6% of the participants were female. The relatively high number of drop-outs reflects two factors: (1) the SMDT was administered close to the end of the online test battery which took between 50 and 120 min to complete; (2) the SMDT required multi-media software to be installed and function on the respondent's computer, and this was not possible for some participants.

Pilot testing of longer versions of the Melody and Rhythm subtests, for item selection purposes, were performed on a smaller sample ($n = 49$; 36 females), mainly consisting of students (age 27.8 ± 9.1 years; mean ± SD) recruited through the website Studentkaninen (www.studentkaninen.se) – a Swedish website for research volunteers.

### 2.2. Materials

The SMDT is composed of three subtests: Melody (18 items), Rhythm (18 items), and Pitch (27 items). In all test items, the task of the participant was to discriminate between two consecutively presented stimuli. Each subtest is constructed so that items become progressively more difficult. Total test taking times for the three subtests were 4.6 ± 1.1 (Melody), 3.2 ± 1.2 (Rhythm), and 3.0 ± 1.3 minutes (Pitch).

#### 2.2.1. Melody

Stimuli in this subtest consisted of isochronous sequences of piano tones. The piano tones were taken from the Kontakt sound library (Steinberg AG). The pitches of the tones ranged from C4 to A#5 (American Standard Pitch; 262–932 Hz). The time interval between tones in a stimulus sequence was always 650 ms. The number of tones per stimulus increased from four to nine as the subtest progressed. For each of these six stimulus lengths, there were three items. Detailed information on the construction of the tone sequences and the selection of the final set of items is provided in the next section. The two stimuli of each item were separated by 1.3 s of silence. The pitch of one randomly selected tone was always different in the second stimulus as compared to the first stimulus. Examples of stimulus pairs for items with a stimulus length of four and nine tones are given in Fig. 1A and B. The sequence was graphically depicted as a straight horizontal line of dots which changed colour when the corresponding tone was played. The task of the participant was to indicate which tone in the second melody was different from the first. Responses were given either by pressing the computer key corresponding to the ordinal number of the differing note, or by clicking on the corresponding dot with the mouse pointer.

#### 2.2.2. Rhythm

In Rhythm, stimuli consisted of rhythmic sequences of brief sine tones. The sine tones were 500 Hz sine waves with a total duration of 60 ms. The loudness of the tone was constant during the first 30 ms and then decreased linearly to 0 db. The inter-onset intervals between tones within a stimulus sequence had a duration of 150, 300, 450, or 600 ms. The number of sounds in each stimulus increased from five to seven as the subtest progressed, with six items for each number of sounds. The two stimuli of each item were separated by 1 s of silence. In 11 out of the 18 Rhythm item the two stimuli differed. In the remaining seven Rhythm items the two stimuli were identical. Further details on the construction of the stimulus sequences and the selection of the final set of items are provided in the next section. Examples of stimulus pairs for items using different stimuli with a sequence length of five and seven notes, respectively, are shown in Fig. 1C and D. The task of the participant was to judge whether the two stimuli were the same or different. Responses were given by pressing either one of two keys on the keyboard or by clicking one of two icons with the mouse pointer.

#### 2.2.3. Pitch

In the Pitch sub-test, stimuli consisted of sine tones. Each tone had total duration of 590 ms and started with a 30 ms ramp from