

Validation of the abbreviated version of the Profile of Musical Perception Skills (Mini-PROMS) in a Spanish sample of musicians and non-musicians

Laura Vázquez-Fragua

Universidad Complutense de Madrid, Spain

Miguel Ángel Fernández-Blázquez

Universidad Autónoma de Madrid, Spain

José María Ruiz-Sánchez de León 

Universidad Complutense de Madrid, Spain

Musicae Scientiae

2024, Vol. 28(4) 619–633

© The Author(s) 2024

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/10298649241227812

journals.sagepub.com/home/msx



Abstract

Despite the efforts made throughout the 20th century to develop standardized batteries to measure musical ability, there seems to be no consensus as to the construct's gold standard. The Profile of Music Perception Skills (PROMS) was created with the aim of overcoming some of the limitations of previous assessment tests. The Mini-PROMS is the shortened version of the full battery and has a shorter application time. The aim of the present study was to validate the Mini-PROMS battery with a sample of Spanish musicians and non-musicians in order to provide useful scales for the development of research in the field of musical skills in the Spanish population. The results show that Mini-PROMS in Spanish is a reliable and valid tool for measuring the musical ability of musicians and non-musicians. Furthermore, this study provides information on some of the differences between musicians and non-musicians, offering useful scales for comparing the results of individuals with those of their reference population.

Keywords

music, cognition, cognitive test, psychological assessment, psychometrics, validation study

The term *musical ability* is used to describe sensitivity to music, the ability to understand music, and/or the ability to produce music. Researchers generally adopt one of two approaches to describing and measuring musical ability. One approach is to categorize individuals in two groups, as musicians and nonmusicians, and compare the performance of the two groups on

Corresponding author:

José María Ruiz-Sánchez de León, Experimental Psychology Department, Facultad de Psicología, Universidad Complutense de Madrid, Campus de Somosaguas, Pozuelo de Alarcón, Madrid 28223, Spain.

Email: jm.ruiz.sdl@psi.ucm.es

an outcome of interest. When musicians outperform nonmusicians, the result is interpreted as evidence for a link between musical ability and the outcome being studied, such as working memory, spatial ability, or language processing skills. A limitation of this approach is that potentially significant gradations in musical ability among musicians and nonmusicians cannot be identified. Nonmusicians may be musical, but their talent may have remained undiscovered, or circumstances may have prevented its development (PROMS; Law & Zentner, 2012). The other approach is to measure individuals' musical abilities using tests of aural ability (*ear tests*). An advantage of this approach is that such tests can be used to identify gradations in the musical skills of individuals who have not received training in music. The musical experience of the listener is at the heart of all musical activity (Sloboda, 2000). From this perspective, we can define music ability as the ability to perceive musical stimuli such as small changes in pitch, loudness, rhythm, and other sub-domains of music processing (Faßhauer et al., 2015).

One of the first researchers to try to measure musical ability using ear tests was Carl Stumpf, who asked students to carry out a range of tasks in which they were required to discriminate small changes in pitch, rhythm, and loudness in stimuli that differed in one sub-domain of music such as rhythm, melody, or accent (Stumpf, 1890). Stumpf also compared the test results with students' own reports of their musical ability to find out if these correlated with his own measures of musical ability. Stumpf's approach was purely exploratory and did not lead to the creation of a test battery designed to measure musical ability objectively. This was not achieved for the first time until nearly 30 years later, when Seashore (1919) published *The Psychology of Musical Talent*, followed by the tests produced by Schoen (1923), Lowery (1926), Dykema (1927), and Seashore (1930). All the tasks in these tests were also based on discrimination between two stimuli that could differ in one musical sub-domain such as rhythm, tone, and pitch. Finally, the Drake Test of Musical Talent (Drake, 1933) sub-tests were found to be correlated with those of Seashore (1930) and Kwalwasser and Dykema (1930), so this test was used by early researchers interested in the neurobiological substrates of musical ability (Grede et al., 1978; Martin, 1964). Years later, the first test designed to assess musical ability was standardized: the Seashore Measures of Musical Talents (Seashore et al., 1956). This test considered not only aural recognition but also other parameters of musicality including auditory, motor, associative, affective, and emotional responses. This battery was widely used and even revised at the end of the 20th century (Seashore et al., 1992). In fact, it was for many years the only tool in Spanish available to assess the construct of musical ability.

Until this time musical ability had been considered innate, present only in individuals with special musical skills (Ericsson et al., 1993; Howe et al., 1998). However, Gordon (1965) suggested that while musical skills could be considered as innate abilities, they could also be taught or modified during the course of development through practice. He developed various tests to measure them at each stage of development: the Music Aptitude Profile (MAP; Gordon, 1967), Primary Measures of Music Audiation (PMMA; Gordon, 1986b), Intermediate Measure of Music Audiation (IMMA; Gordon, 1986a), and Advanced Measures of Music Audition (AMMA; Gordon, 1984). Subsequent authors who developed other tests of musical abilities used these tests to ensure convergent validity through correlational analyses (McLeish, 1968; Rizzo & Vispoel, 1992; Wing, 1962; Young, 1972, 1973);

More recently, Peretz and Coltheart (2003) designed a model of music processing based on studies of brain-damaged patients. This model consisted of temporal and melodic processors working together to send information to a musical lexicon that formed the musical repertoire.

An alteration in any of the connections between the processors could explain the difficulties in musical integration suffered by patients with brain damage. To evaluate the deficits that occur in amusia, they developed the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003), which included six subtests evaluating different sub-domains of music processing: melody, intervals, scale types, rhythm, meter, and musical memory. The MBEA is useful for assessing clinical populations such as people with congenital (Pfeifer & Hamann, 2015) or acquired amusia (Sihvonen et al., 2017), but not for assessing the musical skills of professional musicians.

Thus, the measurement of musical ability using ear tests to evaluate different sub-domains of music processing in this first quarter of the 21st century has been shown to have some limitations: on one hand, (1) the classical tests developed during the 20th century presented some methodological problems in their construction; and (2) did not have the psychometric properties required today; in addition, (3) the sound samples used in these tests were inaccurate as they involved recordings of live performances into which musicians may have introduced small errors; or (4) the recordings were of poor quality due to the recording and preservation methods used; and, on the other hand, (5) the most recent tests are very much aimed at clinical populations, so they are not sensitive to individual differences between healthy adults. To overcome these limitations, Law and Zentner (2012) created the Profile of Music Perception Skills (PROMS). This battery measures perceptual musical skills objectively across nine different music sub-domains (melody, pitch, tone, tuning, rhythm, embedded rhythm, metric accent, tempo, and loudness); was developed using modern digital audio samples; and its reliability and validity were assessed in psychometric studies. In each subtest, according to the instructions, the participant listens twice to a standard stimulus, followed by a comparison stimulus, and is asked to judge whether the comparison stimulus is the same as or different from the standard stimulus (<https://musemap.org/resources/proms>). The test is equally suitable for listeners with all levels and types of musical training, so it is more comprehensive than previous batteries in terms of the musical sub-domains assessed, and it aims to measure each component of each skill with maximum specificity. Moreover, the test items do not relate to any linguistic or cultural aspect characteristic of any particular language, it is available online with instructions in several languages, and it can be used free of charge via the website of the Laboratory for Personality, Emotion and Music of the Faculty of Psychology at the University of Innsbruck (<https://musemap.org/resources/proms>).

The test has adequate psychometric properties for the composite score, with indicators of internal consistency and test-retest reliability, $r < .85$. Convergent validity is established with the melody, rhythm, and tempo dimensions of Gordon's (1967, 1984) MAP and AMMA, the rhythm dimension of the Musical Ear Test (Wallentin et al., 2010), and the timbre dimension in instrumental sample sounds; and there are significant relationships between test performance and external indicators of musical competence ($.38 < r < .62$, $p < .01$). In addition, a divergent validity study was performed, showing the lack of correlations between test scores and a non-musical auditory discrimination task ($r = -.05$, NS) (Law & Zentner, 2012).

It takes 60 min to complete the PROMS, which can cause some difficulties when it is administered. For this reason, two short versions have been created, the Mini-PROMS (Zentner & Strauss, 2017) and the Micro-PROMS (Strauss et al., 2023). The aim of the present study was to evaluate the validity of the Spanish Mini-PROMS with a sample of Spanish musicians and non-musicians to provide useful scales for the development of research on musical skills in the Spanish population.

Method

Participants

Zhang et al. (2020) reviewed 730 papers published between 2011 and 2017 to establish a consensus for the use of the term *musician*. They concluded that there is a six-year rule such that, to be defined as a musician, an individual must have spent at least six years developing expertise in music through training and practice.

Our sample consisted of 230 participants aged between 18 and 47 years ($M=21.46$, $SD=4.11$). A total of 121 (52.6%) met the criterion to be defined as musicians, as they were students at a Superior Conservatory of Music (aged between 18 and 47 years ($M=22.75$, $SD=4.47$), and 109 (47.4%) were controls, as they were students in disciplines unrelated to music at the Complutense University of Madrid.

The distinction between musicians and controls was based on years and intensity of practice. Students at a Superior Conservatory of Music are likely to have studied for at least 10 years before starting their studies in higher education, and to have practiced daily for between 4 and 8 hr for at least 3 years. Our sample were all classical musicians. Although there were some participants in the control group who played or sang music as a hobby, they did not meet Zhang et al.'s criterion for being defined as musicians.

The socio-demographic characteristics of the two groups are shown in Table 1. The musicians were, on average, 2 years older than the controls, and had had more years of education. They also had a more homogeneous gender distribution.

The participants assigned to the group of musicians were students at the Conservatorio Superior de Castilla y León, Conservatorio Superior de Música de Málaga, Conservatorio Superior de Música de A Coruña, Conservatorio Superior de Música de Alicante, and Conservatorio Superior de Música Navarra. They were recruited through information posters and via social networks at each institution. The assessments were conducted individually at each institution between June 1, 2021, and April 27, 2022, in a soundproof booth. The participants assigned to the control group were students taking Bachelor of Psychology and Bachelor of Speech Therapy degrees in the Faculty of Psychology at the Complutense University of Madrid, where information was provided about the research project. The assessment sessions were conducted individually between April 21, 2021, and December 17, 2021, in a soundproof booth at the same faculty. In all cases, appointments were scheduled and managed via a digital platform designed for appointment management.

Instrument

The Spanish Mini-PROMS (Zentner & Strauss, 2017), one of the two short versions of the full PROMS, was administered according to the following instructions:

In each subtest, the participant listens twice to a standard stimulus, followed by a comparison stimulus, and is asked to judge whether the comparison stimulus is the same as or different from the standard stimulus. The Melody subtest assesses the participant's ability to recognize whether or not two short melodies are the same. The Tuning subtest assesses the ability to recognize if one of the notes of a chord in the comparison stimulus is mistuned. The Accent subtest assesses the ability to recognize whether or not two short series of accentuated rhythmic clicks are the same, and the Tempo subtest assesses the ability to recognize whether a synthetic rhythmic structure or a recorded sample of music is being played for the second time faster or slower. (<https://musemap.org/resources/proms>)

Table 1. Socio-demographic characteristics of the sample.

	Total (N= 230)	Controls (n= 109)	Musicians (n= 121)	Sig. ^a	Effect size ^b (95% CI)
Age (mean [SD])	21.46 (4.11)	20.02 (3.10)	22.75 (4.47)	<.001	0.33 [0.21, 0.44]
Years of schooling (mean [SD])	14.84 (2.62)	13.61 (1.78)	15.96 (2.75)	<.001	10.45 [0.34, 0.55]
Gender (n [%])					
Woman	173 (75.22)	92 (84.40)	81 (66.94)	.002	0.2 [0.07, 0.33]
Male	57 (2.478)	17 (15.60)	40 (33.06)		

Note: CI: confidence interval; SD: Standard deviation.

^aStatistics: Mann-Whitney test for numerical variables and Fisher's test for double-entry tables.

^bCorrelation coefficient (*r*).

Test–retest reliability of the Mini-PROMS total score was $r = .87$. Internal consistency reliability was $\omega = 0.87$. The association between the Mini-PROMS total score and the musical background composite index was $r = .52$ ($p < .001$).

Procedure

An ad hoc interview was conducted to collect socio-demographic data and musical education, followed by administration of the Mini-PROMS. The test was conducted in its digitized version, hosted on the Lime Survey platform using an iPad Air and professional Yamaha MT7 headphones. The procedure guaranteed anonymity and the impossibility of linking each participant to their answers.

In all cases, participants gave their explicit consent to participate before they were enrolled in the study by ticking the box “I am 18 years of age or older and I consent to my voluntary participation in this study.” All information gathered in the study was treated and stored in accordance with Spanish legislation. Anonymity was guaranteed throughout the process by codifying responses to prevent subsequent attribution to individuals.

Statistical analyses

The information collected from the questionnaires was entered into a spreadsheet for further analysis using the statistical program R version 4.0 (R Core Team, 2021). Statistical significance was obtained with p values $< .05$. An initial scan of the database was performed to identify the possible presence of missing values and outliers. A preliminary descriptive analysis was then performed using the non-parametric Fisher’s test for double-entry tables, and Mann-Whitney tests for numerical variables, to determine whether there were significant differences between the two groups (musicians vs. controls) in terms of socio-demographic characteristics and Mini-PROMS scores. Effect size was examined using the correlation coefficient (r) statistic. McDonald’s omega (ω) was calculated to assess the internal consistency of Mini-PROMS and its subscales. This is similar to Cronbach’s alpha but provides a more accurate estimate of reliability (Revelle & Zinbarg, 2009; Zinbarg et al., 2006). Next, a multivariate study of the five Mini-PROMS subscales was performed to determine the best way to obtain normative data for the Mini-PROMS. Their distribution was examined to check whether they met the assumptions of statistical normality and homogeneity of variances. Because none of them met these assumptions, non-parametric generalized linear models (GLM) with Poisson distribution and logit link function were used to explore the possible effect of group on Mini-PROMS scores. Age, years of education, and gender were used as covariates to fit the models. Additionally, a logistic regression analysis to examine the relationship between the four subscales of the Mini-PROMS and musicianship was performed. For this, we used the group as the dependent variable, with two levels (musicians vs. controls), and included all four predictor subscales of the Mini-PROMS in the model.

Normative data for the Mini-PROMS were obtained subsequently. Based on the GLM results, it was considered most appropriate to find the normative data for the two groups separately without correction for socio-demographic variables. The direct scores were first assigned to percentile ranks according to their placement in the distribution. The percentile ranks were then converted to scalar values to achieve a Gaussian distribution of values. Finally, a polytomous Rasch model was used exploratively to analyze the behavior of the different items of the Mini-PROMS and to examine possible differences between the two groups. The full database can be downloaded from <https://doi.org/10.6084/m9.figshare.21917301.v1>.

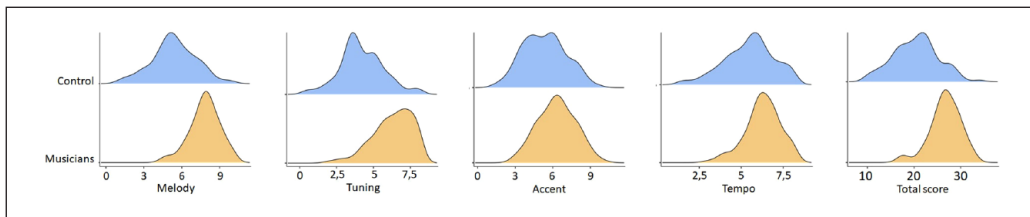
Table 2. Scores on the different subscales of the Mini-PROMS.

	Total (N = 230)	Musicians (n = 121)	Controls (n = 109)	Sig. ^a	Effect size ^b [95% CI]
Melody (mean [SD])	6.70 (1.89)	7.81 (1.19)	5.46 (1.75)	<.001	0.6 [0.52, 0.70]
Tuning (mean [SD])	5.01 (1.64)	5.94 (1.26)	3.97 (1.38)	<.001	0.60 [0.50, 0.68]
Accent (mean [SD])	5.85 (1.66)	6.31 (1.49)	5.34 (1.70)	<.001	0.29 [0.17, 0.41]
Tempo (mean [SD])	5.89 (1.34)	6.28 (1.04)	5.46 (1.49)	<.001	0.31 [0.18, 0.42]
Total score (mean [SD])	23.40 (5.02)	26.34 (3.22)	20.23 (4.69)	<.001	0.61 [0.51, 0.69]

Note: CI: confidence interval; SD: Standard deviation.

^aStatistics: Mann-Whitney test for numerical variables and Fisher's test for double-entry tables.

^bCorrelation coefficient (*r*).

**Figure 1.** Density function distribution of musicians and control group's scores.

Note: Density functions for the two groups from the scores obtained on the subscales and the total scores on the Mini-PROMS. Visual inspection of the density functions of the two groups (musicians: blue, controls: orange) reveals differences between the scores of the two groups, particularly on the melody and intonation subscales, and total scores.

Results

As shown in Table 2, there were statistically significant differences between the two groups on all variables representing the different subscales of the Mini-PROMS such that musicians scored higher than controls, with moderate effect sizes for accent and tempo and high effect sizes for melody, tuning, and total score.

Visual examination of the density functions of all Mini-PROMS subscales, shown in Figure 1, revealed significant differences between the two groups, so it was decided to obtain normative data for each group independently of the other.

In terms of internal consistency, the result of the McDonald's test for the Mini-PROMS was $\omega = .84$. The results for each subscale were $\omega = .75$ (melody), $\omega = .73$ (tuning), $\omega = .52$ (accent), and $\omega = .54$ (tempo).

Five regression studies, one for each subscale, were performed to quantify the effect of group (musicians vs. controls), controlling for participants' age, years of education, and gender. As shown in Table 3, the regression models satisfied the assumptions of the Poisson models, showing in all cases a good fit of the data ($G^2 > .05$), no collinearity between predictors (all < 5), and reasonable dispersion of the observed results compared to the values predicted from the Poisson distribution ($\phi < 1$).

The estimated values in the five models showed that musicians scored significantly higher than controls on the melody, tuning, and accent subscales, and total score for the Mini-PROMS ($p < .05$ in all cases), regardless of their sociodemographic characteristics. Although the difference between the scores of the two groups on the tempo subscale was not significant, there was a trend in favor of the musicians ($p = .054$).

Table 3. Generalized linear models for the subscales of the Mini-PROMS.

Model 1: Melody					
Variables	β	Standard error	OR [95% CI]	<i>z</i>	Sig.
Intercept	1.88	.03	6.56 [6.18, 6.96]	62.10	< .001
Group (control vs musicians)	.36	.06	1.44 [1.28, 1.61]	6.17	< .001
Age	-.01	.01	0.99 [0.98, 1.01]	-.35	.725
Years of schooling	.01	.01	1.01 [0.98, 1.03]	.04	.970
Gender	.02	.06	1.02 [0.91, 1.14]	.32	.753
Deviance (G^2) = .99 Collinearity: Group (control vs musicians) = 1.25; Age = 1.76; Years of schooling = 1.95; Gender = 1.06 Dispersion (ϕ) = .39.					
Model 2: Tuning					
Variables	β	Standard error	OR [95% CI]	<i>z</i>	Sig.
Intercept	1.58	.04	4.84 [4.51, 5.18]	44.57	< .001
Group (control vs musicians)	.41	.07	1.50 [1.31, 1.71]	5.94	< .001
Age	.02	.01	0.98 [0.96, 1.01]	1.63	.103
Years of schooling	.02	.02	1.02 [0.99, 1.05]	1.16	.247
Gender	.01	.07	0.99 [0.87, 1.14]	0.03	.972
Deviance (G^2) = .99 Collinearity: Group (control vs musicians) = 1.25; Age = 1.92; Years of schooling = 2.11; Gender = 1.05 Dispersion (ϕ) = .39.					
Model 3: Accent					
Variables	β	Standard error	OR [95% CI]	<i>z</i>	Sig.
Intercept	1.76	.03	5.83 [5.47, 6.20]	54.90	< .001
Group (control vs musicians)	.15	.06	1.16 [1.03, 1.32]	2.45	.014
Age	.01	.01	0.99 [0.97, 1.01]	0.92	.356
Years of schooling	.02	.01	1.02 [0.99, 1.05]	1.09	.275
Gender	.01	.06	1.01 [0.89, 1.15]	0.21	.832
Deviance (G^2) = .99 Collinearity: Group (control vs musicians) = 1.27; Age = 1.84; Years of schooling = 2.05; Gender = 1.05 Dispersion (ϕ) = .45.					
Model 4: Tempo					
Variables	β	Standard error	OR [95% CI]	<i>z</i>	Sig.
Intercept	1.77	.03	5.89 [5.53, 6.27]	55.60	< .001
Group (control vs musicians)	0.12	.06	1.13 [0.99, 1.27]	1.93	.054
Age	0.01	.01	0.99 [0.98, 1.02]	0.11	.914
Years of schooling	0.01	.01	1.01 [0.98, 1.04]	0.56	.578
Gender	0.02	0.06	1.02 [0.90, 1.16]	0.34	.733
Deviance (G^2) = .99 Collinearity: Group (control vs musicians) = 1.27; Age = 1.76; Years of schooling = 1.96; Gender = 1.05 Dispersion (ϕ) = .31.					

Table 3. (Continued)

Model 5: Total score					
Variables	β	Typical error	OR [95% CI]	z	Sig.
Intercept	3.14	0.02	23.16 [22.44, 23.90]	195.05	<.001
Group (control vs musicians)	0.26	0.03	1.29 [1.22, 1.37]	8.23	<.001
Age	0.01	0.01	0.99 [0.98, 1.00]	1.46	.144
Years of schooling	0.01	0.01	1.01 [0.99, 1.02]	1.36	.175
Gender	0.01	0.03	1.01 [0.95, 1.08]	0.43	.665
Deviance (G ²) = .99					
Collinearity: Group (control vs musicians) = 1.26; Age = 1.81; Years in school = 2.01; Gender = 1.04					
Dispersion (φ) = .74.					

Note: PROMS: Profile of Music Perception Skills; OR: odds ratio; CI: confidence interval.

Table 4. Binomial logistic regression to distinguish between musicians and controls.

Predictors	β	<i>SE</i>	OR (95% CI)	<i>z</i>	Sig.	
Intercept	-8.90	1.36	1.16e-4 [9.46e-6, 0.01]	-6.55	<.001	
Melody	.87	.16	2.38 [1.74, 3.25]	5.44	<.001	
Tuning	.76	.16	2.15 [1.57, 2.93]	4.81	<.001	
Accent	-.10	.13	0.91 [0.70, 1.17]	-.76	.446	
Tempo	-.02	.17	0.98 [0.70, 1.37]	-.13	.899	

Note: OR: odds ratio; CI: confidence interval.

As can be seen in Table 4, the results of the logistic regression analysis indicate that the model fitted the data well, with a deviance of 178 on 4 degrees of freedom ($p < .001$) and McFadden $R^2 = .44$. Scores on the melody ($OR = 2.38$, $p < .001$) and tuning subscales ($OR = 2.15$, $p < .001$) were significant predictors of being classified as a musician. For each additional point on the melody subscale, the odds of being classified as a musician increased by 138% whereas, for each additional point on the tuning subscale, the odds of being classified as a musician increased by 115%. Scores on the accent ($OR = 0.91$, $p = .446$) and tempo subscales ($OR = 0.98$, $p = .899$) were not significant predictors of being classified as a musician.

None of the socio-demographic variables influenced the results of the Mini-PROMS. Percentile ranks and their corresponding scalar values were calculated for each group without having to correct for sociodemographic characteristics. Tables 5 and 6 show the normative data for each of the two groups.

The ROC curve for the logistic regression model is shown in FIGURE2Figure 2. The AUC value was .91, indicating that the model had excellent predictive performance. The curve showed that sensitivity decreases, and specificity increases, as the threshold for predicting musicianship becomes more stringent (i.e., moving toward the upper-left corner of the plot). The model had a sensitivity of .89 and a specificity of .85 at the optimal threshold, meaning that the model correctly identified 89% of the musicians and 85% of the controls.

Our logistic regression model thus showed good predictive performance for identifying musicians vs. controls. Of the four Mini-PROMS subscales, only scores for melody and tuning were significant, suggesting that both provide clear distinctions between the two groups. By contrast, the scores on the accent and tempo subscales did not allow us to distinguish between musicians and controls.

Table 5. Scalar scores and percentile ranges from the raw scores of the subscales and the total scores for the Mini-PROMS (musicians).

Scalar score	Percentile ranges	Melody raw score	Tuning raw score	Accent raw score	Tempo raw score	Mini-PROMS raw score
2	< 1	0–4	0–2	0–2.5	0–3	0–17
3	1	4.5	2.5	3	3.5	17.5
4	2–3	5	3	3.5	4	18
5	4–6	5.5–6	3.5–4	4	4.5	18.5–22
6	7–12	6.5	4.5	4.5	5	22.5–23
7	13–20	7	5	5	5.5	23.5–24.5
8	21–30	7.5	5.5	5.5	6	25–25.5
9	31–43	8	6	6	6	26–26.5
10	44–56	8	6.5	6.5	6.5	27
11	57–69	8.5	7	7	7	28–28.5
12	70–79	9	7.5	7.5	7	29–29.5
13	80–87	9	8	8	7.5	30
14	88–93	9.5	8	8.5	8	30.5–31
15	94–96	10	8	9	8	31.5–32
16	97–98	10	8	9	8	32.5
17	99	10	8	9.5	8	33
18	> 99	10	8	10	8	33.5–36

Note: PROMS: Profile of Music Perception Skills.

Table 6. Scalar scores and percentile ranges from the raw scores of the subscales and the total scores for the Mini-PROMS (controls).

Scalar score	Percentile ranges	Melody raw score	Tuning raw score	Accent raw score	Tempo raw score	Mini-PROMS raw score
2	< 1	0–1	0	0–1.5	0–1	0–10
3	1	1.5	0.5	2	1.5	10.5
4	2–3	2	1–1.5	2.5	2–2.5	11
5	4–6	2.5	2	3	3	12.5
6	7–12	3	2.5	3.5	3.5	13–14.5
7	13–20	3.5–4	3	4	4	15–16.5
8	21–30	4.5	3.5	4.5	4.5	17–17.5
9	31–43	5	4	5	5	18–19.5
10	44–56	5.5	4.5	5.5	5.5	20–21.5
11	57–69	6	5	6	6	22–22.5
12	70–79	6.5–7	5.5	6.5	6.5	23–24
13	80–87	7.5	6	7–7.5	7	24.5–25.5
14	88–93	8	6.5	8	7.5	26–27.5
15	94–96	8.5	7	8	8	28–28.5
16	97–98	9	7.5	8.5	8	29–29.5
17	99	9.5	8	9	8	30
18	> 99	10	8	9.5–10	8	30.5–36

Note: PROMS: Profile of Music Perception Skills.

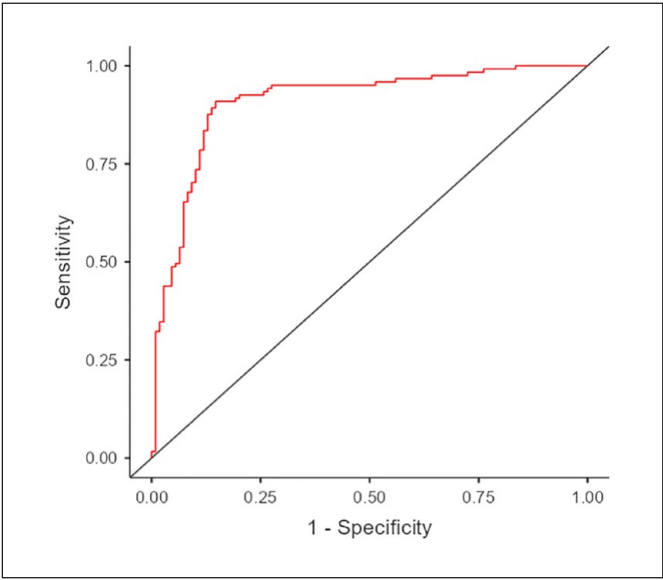


Figure 2. ROC curve for the logistic regression model.

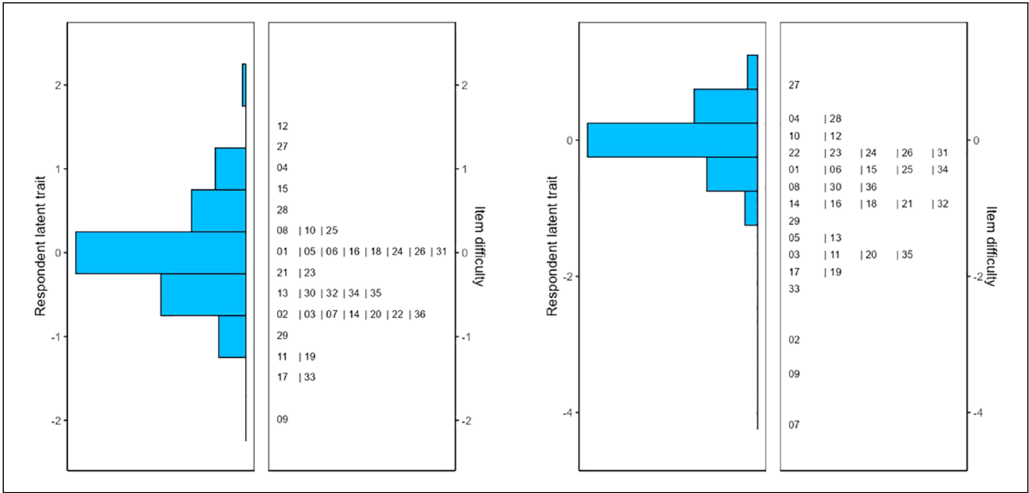


Figure 3. Wright's maps of the Mini-PROMS for the two groups (musicians vs. controls).
Note: The figure shows the relationship between the participants' ability (left) and the difficulty of the items (right) for the control group (left graph) and the group of musicians (right graph).

Finally, a polytomous Rasch model was used to explore the behavior of the items comprising the Mini-PROMS within the groups. Figure 3 shows the ability and difficulty parameters of the items in the two groups musicians vs. controls) using Wright maps. As expected, the musicians were found to score higher on the items. It was also observed, however, that the items were more difficult for controls (range: -1.86 — 1.61) than musicians (range: -4.06 — 0.71), indicating a continuum of musical ability in the control group and a ceiling effect among musicians.

Discussion

Despite the efforts made throughout the 20th century to develop standardized test batteries to measure musical ability, there seems to be no consensus as to a gold standard test for the construct. The main limitations that have been noted with the tests developed over the past century are that they have focused too much on musical perception (Hallam & Shaw, 2002) and that different subtests have often measured a combination of abilities rather than a specific musical ability (Demorest, 1995; Hallam & Shaw, 2002). The PROMS, and hence the Mini-PROMS, aim to find individual differences between healthy adults with and without musical training, making them very useful tools for research in the general population. In addition, the stimuli are carefully selected, balanced, and revised, providing optimal psychometric indicators of reliability and validity. In this sense, both versions overcome the limitations of other previously developed batteries.

The Mini-PROMS has been used widely in recent years, especially when several batteries are administered in one evaluation session (Foncubierta, 2020; Samiotis et al., 2021; Sun, 2022; Talamini et al., 2022). The utility of the battery is also reflected in the number of languages into which it has been translated.

Test–retest reliability of the Mini-PROMS total score was $r = 0.87$. (Zentner & Strauss, 2017). Our contribution is to have evaluated the validity of the Spanish Mini-PROMS; we found a high correlation in form of a group difference in PROMS scores between musicians and non-musicians.

The results of this study show that all the items of the Mini-PROMS are independent of socio-demographic characteristics and the battery can therefore be administered to healthy adults regardless of age or gender. The sociodemographic differences found between the samples correspond to the population from which participants were recruited. According to the latest data (Corrales, 2021), the percentage of women enrolled in Health Sciences is 70.8% and in Arts and Humanities 61.9%, compared to 29.2% and 38.1%, respectively, of men. Thus our sample included more women ($n = 173$) than men ($n = 57$). Moreover, musicians begin their studies at higher education institutions later than university students. This is because these studies are preceded by 10 years of study at elementary and professional music schools, which does not always end at the same time as high school, so it is very common to begin higher education studies in music at the age of 19 or 20 (Corrales, 2021).

Effects of group were seen only for musicians vs. controls; this result is consistent with the expectation that musicians would have more advanced musical skills as the result of their practice. The non-significant effect of group on tempo (although a trend was observed) may be due to the greater familiarity of the general population with this parameter since, unlike melody, accent, and tuning (required for advanced-level music making), it concerns the speed at which the excerpt is played, a common parameter in daily life (e.g., for audio playback), making it easier for non-musicians to spot variations.

Figure 1 shows differences between the density functions for the scores of the two groups on each of the Mini-PROMS subscales, showing less variability in the musicians' scores and thus suggesting that the musicians' group was more homogenous than the control group. Although one of the most interesting attributes of this test is that it can be taken by any individual regardless of their background, musical training, culture, and language, it would be worth considering musicians and non-musicians as distinct and independent populations when assessing individuals' musical abilities. This would provide fuller information about an individual's abilities in comparison with the rest of the population of musicians or non-musicians, as appropriate.

Figure 2 shows that, for controls, both item difficulty and skills were normally distributed; for musicians, only item difficulty was normally distributed. This could be interpreted in two ways: as a ceiling effect on the ability of highly trained musicians to improve their musical skills through practice when they are already performing at a high level, or as a limitation of the battery, which cannot be used to detect subtle differences between such musicians.

One limitation of the study, more widely, is that the control participants were mainly students at a single faculty at a single university. A second limitation is that the original PROMS has not yet been validated with a Spanish sample, so we could not compare the results of the two validations. The original PROMS may not be the first instrument of choice in many research contexts, however, since it takes so long to complete. A third limitation is that we did not include other instruments in our study that would have enabled us to test the Spanish Mini-PROMS for its convergent or discriminant validity, although this would be well worth doing in future.

Nevertheless, we can conclude that the Spanish Mini-PROMS is a reliable and valid tool for measuring the musical abilities of musicians and non-musicians. Furthermore, this study provides information on the differences between the population of musicians and non-musicians, and provides useful scales for comparing an individual's results with those of their reference population. Taking advantage of the opportunities offered by the Mini-PROMS in terms of its quick and easy administration, we propose to conduct multidimensional studies comparing the results obtained from the Mini-PROMS with those obtained from other cognitive and emotional assessments with the aim of increasing our knowledge of the factors underlying musical abilities.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

José María Ruiz Sánchez de León  <https://orcid.org/0000-0003-2834-0827>

References

- Corrales, M.A.P. (2021). Alcance e impacto económico en España del sector cultural y creativo. La Cuenta Satélite de la Cultura. El Anuario de Estadísticas Culturales (Scope and economic impact of the cultural and creative sector in Spain. The Culture Satellite Account. The Cultural Statistics Yearbook). In *Economía y cultura: una mirada hacia el futuro (Economy and culture: a look into the future)* (pp. 21–52). Fundación de las Cajas de Ahorros (FUNCAS).
- Demorest, S. M. (1995). Issues of ecological validity for perceptual research in music. *Psychomusicology: A Journal of Research in Music Cognition*, 14(1–2), 173–181.
- Drake, R. M. (1933). Four new tests of musical talent. *Journal of Applied Psychology*, 17(2), 136–147.
- Dykema, P. W. (1927). Tests and Measurements Department. *Music Supervisors' Journal*, 13(4), 65–70.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406.
- Faßhauer, C., Frese, A., & Evers, S. (2015). Musical ability is associated with enhanced auditory and visual cognitive processing. *BMC Neuroscience*, 16(1), 1–9.
- Foncubierta, J. M. (2020). El proceso lector en español como lengua extranjera: fluidez lectora silenciosa y aptitud musical [The reading process in Spanish as a foreign language: Silent reading fluency and musical aptitude]. *Perfiles de estudiantes universitarios con dificultades lectoras en el aprendizaje del español como lengua extranjera*, 209 <https://shorturl.at/apCFV>.
- Gordon, E. (1965). The Musical Aptitude Profile: A new and unique musical aptitude test battery. *Bulletin of the Council for Research in Music Education*, 6, 12–16.
- Gordon, E. (1967). The Musical Aptitude Profile. *Music Educators Journal*, 53(6), 52–54.

- Gordon, E. (1984). A longitudinal predictive validity study of the Intermediate Measures of Music Audiation. *Bulletin of the Council for Research in Music Education*, 78, 1–23.
- Gordon, E. (1986a). *Intermediate Measures of Music Audiation*. GIA Publications.
- Gordon, E. (1986b). *Primary Measures of Music Audiation and the Intermediate Measures of Music Audiation: Music aptitude tests for kindergarten and first, second, third, and fourth grade children*. GIA Publications.
- Grede, S. E., Parsons, O. A., & Bertera, J. H. (1978). Hemispheric differences in music perception: Aptitude vs. experience. *Neuropsychology*, 16, 369–373.
- Hallam, S., & Shaw, J. (2002). Constructions of musical ability. *Bulletin of the Council for Research in Music Education*, 102–108.
- Howe, M. J., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioural and Brain Sciences*, 21(3), 399–407.
- Kwalwasser, J., & Dykema, P. (1930). Manual of directions for Victor Records. *Carl Fischer*, 302, 63026.
- Law, L. N., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills. *PLOS ONE*, 7(12), Article e52508.
- Lowery, H. (1926). Cadence and phrase tests in music. *British Journal of Psychology*, 17(2), 111.
- Martin, J. G. (1964). Aptitude test score changes following musical training. *The Journal of Educational Research*, 57, 440–442.
- McLeish, J. (1968). *Musical cognition* (Music Education Research Papers 3). Novello & Co.
- Ministerio de Cultura y Deporte (2022). *Anuario de estadísticas culturales 2021* [Cultural Statistics Yearbook 2021]. <https://goo.su/VfSf>
- Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, 999(1), 58–75.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 1, 688–691.
- Pfeifer, J., & Hamann, S. (2015). Revising the diagnosis of congenital amusia with the Montreal Battery of Evaluation of Amusia. *Frontiers of Human Neuroscience*, 9, Article 161.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega and the glb: Comments on Sijsma. *Psychometrika*, 74(1), 145–154.
- Rizzo, T. L., & Vispoel, W. P. (1992). Changing attitudes about teaching students with handicaps. *Adapted Physical Activity Quarterly*, 9(1), 54–63.
- Samiotis, I. P., Qiu, S., Lofi, C., Yang, J., Gadiraju, U., & Bozzon, A. (2021). Exploring the music perception skills of crowd workers. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 9, 108–119.
- Schoen, M. (1923). The validity of tests of musical talent. *Journal of Comparative Psychology*, 3(2), 101.
- Seashore, C. E. (1919). *The psychology of musical talent*. Silver, Burdett & Company.
- Seashore, C. E. (1930). Measures of musical talent: A reply to Dr. C. P. Heinlein. *Psychological Review*, 37(2), 178–183.
- Seashore, C. E., Lewis, D., & Saetveit, J. G. (1956). *Seashore measures of musical talents*. Psychological Corp.
- Seashore, C. E., Lewis, D., & Saetveit, J. G. (1992). *Tests de aptitudes musicales de Seashore (Seashore Measures of Musical Talents): Manual*. Tea.
- Sihvonen, A. J., Ripollés, P., Rodríguez-Fornells, A., Soinila, S., & Särkämö, T. (2017). Revisiting the neural basis of acquired amusia: Lesion patterns and structural changes underlying amusia recovery. *Frontiers in Neuroscience*, 11, Article 426.
- Sloboda, J. A. (2000). Individual differences in music performance. *Trends in Cognitive Sciences*, 4(10), 397–403.
- Strauss, H., Reiche, S., Dick, M., & Zentner, M. (2023). Online assessment of musical ability in 10 minutes: Development and validation of the Micro-PROMS. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-023-02130-4>
- Stumpf, C. (1890). *Tonpsychologie* [Tone psychology]. Hirzel.[BV].
- Sun, L. (2022). *Musical training and improved visual-spatial working memory, verbal working memory and processing speed*. <https://hdl.handle.net/11299/228072>

- Talamini, F., Vigl, J., Doerr, E., Grassi, M., & Carretti, B. (2022). Auditory and visual mental imagery in musicians and non-musicians. *Musicae Scientiae*, 27(2), 428–441.
- Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, 20(3), 188–196.
- Wing, H. D. (1962). A revision of the Wing musical aptitude test. *Journal of Research in Music Education*, 10(1), 39–46.
- Young, W. T. (1972). Musical Aptitude Profile norms for use with college and university nonmusic majors. *Journal of Research in Music Education*, 20(3), 385–390.
- Young, W. T. (1973). The Bentley “Measures of Musical Abilities”: A congruent validity report. *Journal of Research in Music Education*, 21(1), 74–79.
- Zentner, M., & Strauss, H. (2017). Assessing musical ability quickly and objectively: Development and validation of the Short-PROMS and the Mini-PROMS. *Annals of the New York Academy of Sciences*, 1400(1), 33–45.
- Zhang, J. D., Susino, M., McPherson, G. E., & Schubert, E. (2020). The definition of a musician in music psychology: A literature review and the six-year rule. *Psychology of Music*, 48(3), 389–409.
- Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale’s indicators: A comparison of estimators for ω_h . *Applied Psychological Measurement*, 30(2), 121–144.