Original Research Article

National Association for Music Education

Journal of Research in Music Education
2023, Vol. 71(2) 188–207

DOI: 10.1177/00224294221134571
jrme.sagepub.com
SAGE

About Crows and Nightingales: Which Factors Predict Vocal Abilities in Elementary School Children?

Zyxcban G. Wolfs[1], Henny P. A. (Els) Boshuizen[1], and Saskia Brand-Gruwel[1]

Abstract

In this study, we explored associations between perception of musical features (pitch, tonality, timing, timbre, and loudness), vocal abilities (singing accuracy, melodic achievement, and rhythmic performance), and cognitive levels (reading accuracy and comprehension, grammar, and math) among children ages 6 to 12. Musical perception abilities were measured using the Implicit Tonal Ability Test. Vocal abilities were measured using the Vocal Musical Ability Test. Cognitive levels were measured using standardized Dutch performance tests on academic skills. We investigated which factors (age, gender, cognitive levels, school type, music perception abilities, and participation in music education) predict vocal abilities and how these abilities differed by age. Results showed that singing accuracy was best predicted by gender, math level, and music perception abilities. Melodic achievement was best predicted by age, school type, math level, and music perception abilities. Rhythmic performance was best predicted by age, instrumental music education, and music perception abilities. Regardless of their age, differences in singing abilities between children were large. We advise teachers to provide activities in which repetition and prediction of patterns, scales, all intervals, and intonation are practiced, including transposing melodies and repeating rhythmic patterns.

Keywords

vocal abilities, singing accuracy, musical perception, enculturation, rhythmic performance

[1]Open University of the Netherlands, Heerlen, Netherlands

Corresponding Author:
Zyxcban G.Wolfs, Open University of the Netherlands, Valkenburgerweg 177, Heerlen, 6401 DL, Netherlands.
Email: z.wolfs@gmail.com

Perception of music, singing accuracy, the use of singing voice, and melodic vocal achievement are intrinsic parts of music education around the world. Although elementary music curricula vary widely, many include singing songs as a primary content area as well as a method for teaching (Salvador, 2010). While children sing, they face challenges in both vocal performance and memorization with regard to a song's melody, rhythm, tempo, and lyrics. In recent decades, children's singing development and achievement (Philips & Doneski, 2012; Welch, 2015), their voice characteristics (Cuadrado & Rusinek, 2016), and use of singing voice and singing accuracy (Rutkowski, 2019) have been extensively studied with a general focus on how to improve singing instruction and how to deal with those pupils who are monotones. For both music educators and researchers in music cognition, the role of maturation versus training with regard to singing accuracy remains debatable. One reason for the ambiguity is the lack of uniformity with which singing accuracy is defined and measured (Pfordresher & Brown, 2007; Salvador, 2010). Another reason is the apparently reversible nature of singing accuracy development. For example, Demorest and Pfordresher (2015) found that singing accuracy improves considerably in children from kindergarten to late elementary school yet reverses dramatically such that untrained adolescents sing at the level of preschool children.

An essential question was raised by Gould (1969), to which the answer remains partially unsolved: "Could the causes of difficulties in perceiving tones, in remembering melodies and in correlating the vocal mechanism with them be identified?" (p. 12). Since then, many researchers investigated pitch discrimination abilities as a possible source of poor singing and pitch-matching; see Rutkowski (2015) for an overview. Most studies found only weak or no correlations between perception and vocal abilities. In addition to tonal awareness, physical skill development and kinesthetic activities have been mentioned as contributing to singing accuracy (Szabo, 2001). Nevertheless, many studies only focused on perception of pitch and thereby ignored other musical features, such as perception of tonal features (scale, tonal function, harmonic intervals), timing (rhythmic patterns and beat), timbre, and loudness. Although pitch is the most well-known and dominant musical feature in Western musical culture, children are frequently exposed to the other musical features as well. To improve the quality of elementary music curricula, it is important to investigate how musical features other than just pitch relate to children's vocal skills and their general cognitive abilities. Corrigall and Schellenberg (2015) noted that "research on the development of music perception and cognition is still in its infancy, and some areas remain virtually unexplored" (p. 18). In the following, we discuss the perception of musical features, salient concepts within music cognition, and the vocal performance abilities of elementary children.

Musical Features and Musical Enculturation

Children do not require theoretical knowledge about the Western tonal system or musical reading skills to perceive and interpret acoustic signals within a culturally defined tonal-bound system. How do children then, without help from a teacher or teaching aids, acquire implicit knowledge and expectations of cultural achievements in the domain of music? The answer to this question can be found in a process that starts before birth, when fetuses older than 33 weeks already can process complex acoustic signals (Kisilevsky et al., 2004). Immediately after birth, music becomes omnipresent, and while children grow up, they may listen to music for various reasons, being unaware of a remarkable and latent cognitive process that unfolds: the enculturation of their minds to a culture-bound tonal system (Corrigall & Trainor, 2014; Matsunaga et al., 2015). Music enculturation is the natural development of music schemata through the shaping influences of the environment, which takes place from a very young age and can continue over the lifetime (Morrison et al., 2008). Most research suggests, however, that children do not become enculturated listeners until they reach the age of 3 or 4 (Corrigall & Schellenberg, 2015). In combination with general cognitive development, perception of musical features function as building blocks for a first step toward musical cognition. Pearce and Rohrmeier (2012) distinguished between pitch, tonality, timbre, timing, stress, loudness, and spatial location as self-contained musical features. In this study, we refer to these perception skills as "implicit tonal abilities" because no explicitly taught knowledge is required.

Children's Vocal and Rhythmic Abilities

The question of whether and how tonal abilities are related to vocal performance abilities has been mainly investigated in terms of pitch. Many studies confirmed a gradual increase in vocal range and pitch-matching accuracy through maturation (for an overview see Welch, 2015). For young children, these two task types showed a moderate positive relationship (Demorest et al., 2018). Although the vocal apparatus is similar until puberty, some studies reported differences in vocal performance skills with regard to gender; see Hedden (2012) for a review. With regard to pitch-matching tasks, researchers have mainly focused on the role of age and gender. Pitch-matching appears to improve with age, is dependent on the singer's vocal range, and requires tonal memory while singing. For young boys and girls, pitch-matching results tend to be similar, although some studies have indicated that boys sing less accurately as they grow older (e.g., Mang, 2006; Welch et al., 2012). Furthermore, Pfordresher and Demorest (2021) found significant correlations between age as well as musical training and singing accuracy in participants ages 6 to 99.

Very little research has been done on children's performance abilities and knowledge about grouping structure (Corrigall & Schellenberg, 2015). Drake and Gérard (1989) concluded that children ages 5 to 7 are better in reproducing regular rhythmic patterns rather than irregular patterns. Although perception of musical beat is already present in infants, synchronizing to the beat takes much longer to develop (Bergeson & Trehub, 2006). Children prefer a faster tempo than adults (McAuley et al., 2006).

Music Cognition in Childhood

In addition to the enculturation process and maturation of vocal mechanisms, cognitive abilities are required to give meaning to music. Pearce and Rohrmeier (2012) described the comprehensiveness of musical cognition when they stated that

> musical listening, performance, and interaction involve a wide range of cognitive functions and processes, including auditory scene analysis, streaming, attention, learning and memory, formation of expectations, multimodal integration, recognition, syntactic processing, processing of forms of meaning, emotion, and social cognition. (p. 473)

From an educational perspective, Gordon (2012) distinguished between audiation, imitation, recognition, comprehension, and memorization as conditional cognitive concepts for understanding musical structures. Therefore, Gordon recommends a music curriculum for children that offers and trains these concepts.

In the curriculum of the vast majority of elementary schools, however, the focus of cognitive development is on teaching academic skills. This concerns skills in mathematics, grammar, reading comprehension, and reading accuracy. In research into the relationship between cognition and musical skills, it is common to include academic skills as well. In this study, we refer to early academic skills as "cognitive levels."

Assessment of Musical Abilities

Measuring musical abilities in elementary education is a complex matter. Besides the challenges posed by the wide variety in functions, processes, and concepts, children's limited cognitive abilities have to be taken into account as well. Smaller memory capacities, lower comprehension skills, and shorter attention spans can make it cumbersome to obtain reliable measures about children's actual knowledge and skills with one instrument that covers the full age range of elementary school. Corrigal and Schellenberg (2015) therefore advise to design engaging and child-friendly tasks with simple instructions, a limited number of response options (e.g., two alternatives), and a limited number of trials.

Implicit tonal abilities and cognitive abilities can be measured by taking listening tests and academic achievement tests if these are appropriately designed in terms of content, procedure, and appearance for children in elementary education. However, assessing untrained children's singing accuracy remains a challenge (Nichols, 2017). Existing literature reports different age ranges, various difficulty levels, different definitions of accurate singing, and various types of measures used for evaluation and scoring. Apart from challenges with regard to reliability and validity, the inability to sing can have various reasons of its own. According to Corrigall and Schellenberg (2015), cognitive constraints (memory limitations, lack of knowledge), difficulty with regard to controlling the motor movements of the vocal apparatus, or a lack of motivation to perform well can cause poor singing. As a result, music teachers also face challenges in assessing singing achievements (Salvador, 2010). Over the years, several measurement instruments were developed (Rutkowski, 2019), often without reports for reliability and validity measures; see Salvador (2010) for a critical review on this matter. Nonetheless, researchers have provided elementary music educators with measures of singing achievement, such as how to assess children's vocal range and singing accuracy (Hedden, 2012; Rutkowski, 2019; Salvador, 2010). According to Nichols (2016a), two distinct task types can be distinguished in tests for assessing singing accuracy: pitch-matching (repeat individual pitches, intervals, and patterns) and song material (sing phrases or entire songs). Nichols (2016b) reported that children ages 6 to 11 were better in reproducing single intervals and pitches than reproducing four-note patterns and song singing. The most efficient task for assessment purposes is having children echo tonal patterns (Rutkwoski, 2019). Although minor and descending patterns and skips are easier to perform for children, patterns encompassing an octave are recommended to be included in such assessments.

In summary, generalizing among results remains rather difficult due to a lack of consistency in approaches of assessment. Besides, a narrow focus on, for example, just pitch or singing accuracy makes existing tests less suitable for research with a broader scope. Furthermore, development of appropriate and reliable tasks for various age ranges is a time-consuming process.

Research Questions

Musical enculturation, maturation, and learning effects might contribute to the musical development of children. Consequently, the speed and quality of musical development may depend on characteristics of the learning environment and children's innate traits. Elementary music education is usually taught to children based on their grade and age instead of their skills. Extant literature points to differences in vocal skills based on, for example, age and gender (Hedden, 2012; Mang, 2006; Pfordresher & Demorest, 2021; Welch et al., 2012). However, the development of children's vocal abilities is not understood entirely (Corrigall & Schellenberg, 2015). Little is known about musical perception abilities, cognitive academic levels, and their potential relation with vocal abilities. The structure of the curriculum as well as the instruction of teachers to their students would benefit from gaining more insight into whether and how musical abilities are related to other characteristics. This would enable educators to craft better music curriculums and instruction. To this end, we pose the following research questions: (1) Which factors (age, gender, cognitive levels, school type, music perception abilities, and participation in music education) predict vocal abilities (singing accuracy, melodic achievement, and rhythmic performance)? and (2) Do children of different ages perform differently with regard to subcategories of vocal abilities, such as transposing melodies, reproducing rhythmic patterns, beat-keeping, completing melodic patterns, and audiating tonal distances?

Method

Participants

Data were collected from 233 elementary school children (122 females) between ages 6 and 12: 6 ($n = 37$), 7 ($n = 37$), 8 ($n = 38$), 9 ($n = 39$), 10 ($n = 44$), and 11–12 ($n = 38$, of which $n = 6$ for the age of 12). This age range was selected because (a) different levels of musical enculturation and singing accuracy have been reported in this age range (Nichols, 2016b; Speer & Meeks, 1985; Trainor & Trehub, 1992) and (b) many children start formal music lessons in this age range. Participants received education at two elementary schools in the south of the Netherlands (school type: a Waldorf school located in a city, n = 187, and a Dutch Public school located in a rural area, n = 46). Some elementary reading skills were required in order to be able to read and understand test items. Reading music notation was, however, not required because none of the musical test items relied on visual input. Informed consent was obtained from the parents of all participants. Demographic information was gathered with regard to age, gender, school type (Waldorf school vs. public school), and participation in music education outside elementary school (no music education, instrumental education, vocal education, or instrumental and vocal music education).

Measurement Instruments

Vocal Musical Ability Test. The Vocal Musical Ability Test (VMAT) consisted of three components of vocal abilities: singing accuracy (20 items), melodic achievement (22 items), and rhythmic performance abilities (15 items) and took 20 minutes for each participant to complete (see Appendix SA in the supplemental materials included with the online version of this article for a comprehensive overview). The items were derived from the Sung Performance Battery by Berkowska and Dalla Bella (2013) and the AIRS Test Battery of Singing Skills (see Cohen, 2015) for measuring vocal abilities. Because these tests mainly cover pitch-matching, singing melodies, and improvising, items measuring more advanced vocal and cognitive musical abilities were added, such as transposing, audiating tonal patterns, and pattern completion. The difficulty level of the items was adjusted to the target group in terms of vocal range, test length, and procedure.

To measure singing accuracy, it is important to assess children's use of singing voice first. After all, children who sing monotonically will also score low on singing accuracy. Our test included several pitch-matching tasks, such as single pitch, interval pitch, and patterns, as well as song-singing because these tasks are considered to be adequate discriminators of singing accuracy (Nichols, 2016b; Roberts & Davies, 1975). Children's voices were doubled by another human voice because existing research showed that vocal doubling yields better singing accuracy performance than piano doubling (Nichols, 2020). However, not all children who do use their singing voice will necessarily sing in tune. Moreover, for sufficient reliability and validity of results, it is essential to define "the line between lacking precise intonation and an incorrect pitch" (Salvador, 2010, p. 43). With the VMAT, we tried to take these issues into account. By rigorously assessing different pitch-matching tasks for intonation errors, we obtained measures for singing accuracy.

Melodic achievement tasks were assessed for correct pitches. Tasks were not assessed for minimal intonation deviations because these were focused on unraveling cognitive processes, such as imitation, recognition, memorization, audiation, and comprehension. However, any absolute difference greater than approximately 50 cents was considered an error (where 100 cents = 1 semitone). Examples of melodic achievement tasks are transposing a melody, audiating specific tonal distances, and completing melodic patterns.

Rhythmic performance abilities were assessed for the correct tempo and execution of correct rhythmic patterns. The tempo was considered correct if the correct beat was held for 90% of the duration of an item. Rhythmic patterns were considered correct if no deviations greater than the value of a 32nd note in the given tempo were present, with the exception for notes or rests greater than a quarter, for which a deviation of a 16th note was allowed.

All performance tasks were assessed in terms of pitch and timing errors. The first author therefore analyzed each result by ear, and a second coder scored 25% ($n = 58$) of the audio files. Cohen's κ values ranged from 0.878 to 1.000. All κ values and standard errors are displayed in online Table S1, as well as the number of items for each VMAT component. There were no missing data, apart from two participants who only completed rhythmic tasks.

Implicit Tonal Ability Test. The Implicit Tonal Ability Test (ITAT) was developed and consisted of 49 multiple-choice test items (3 to 7 items per musical feature) regarding implicit tonal abilities with four possible answers (including an "I don't know" option). No music reading skills or music theoretical knowledge were required to take the test. The Swedish Musical Discrimination Test (Ullén et al., 2014) for measuring musical perception abilities and the Montreal Battery of Evaluation of Musical Abilities (Peretz et al., 2013) were used as a starting point for item development; items focusing on noticing alterations, idiom of consonance, tonal distances in harmonic intervals, and differences in timbre and loudness were added. For each item, a short audio clip was created with a length between 10 and 30 seconds (see online Figure S1 for example items). Each subscale focused on sensitivity toward one key feature of the Western tonal system, specifically, pitch and intonation (microtones), tonality (tonal function, scales, and intervals), timing (rhythm and beat), timbre, and loudness. Cronbach α coefficients for item sets belonging to ITAT subscales were .627 for pitch, .286 for intonation, .256 for tonal function, .420 for scales, .450 for intervals, .096 for timing, .442 for timbre, and .411 for loudness. The ITAT subscale constructs did not operate reliably, but the one-dimensional composite scale score did ($\alpha = .70$). Rasch model-based analysis of dimensionality and differential item functioning showed that the ITAT discriminated between high and low performers and that all ITAT items behaved as one Rasch dimension in the test. Reliability measures were sufficient for persons ($\alpha = .76$) and items ($\alpha = .98$).

Standardized school performance tests. Dutch elementary school exams (Central Institute of Test Development, 2021) on mathematics, reading comprehension skills, reading accuracy, and grammar in elementary education were used to obtain scores for participants' cognitive levels. Individual scores, also known as proficiency scores, are determined by comparing them with the Cito national database, which for this purpose is divided into 5 percentile categories expressed by Roman numerals (I = highest 20%, V = lowest 20%; see online Table S2). Simulated test–retest values based on 1 million entries ranged from 0.86 to 0.96. Validation reports per grade level and subject show the validity and reliability measures of each test (Hop et al., 2016, 2017, 2019; Hop & Engelen, 2017; Janssen et al., 2015; Jolink et al., 2015a, 2015b; Tomesen et al., 2015a, 2015b; Tomesen, Weekers, Hilte, Jolink, & Engelen, 2016; Tomesen, Wouda, & Horsels, 2016; Tomesen, Weekers, Hiddink, & Jolink, 2017; Tomesen, Wouda, & Horsels, 2017; Tomesen, Engelen, & Hiddink, 2018; Tomesen, Wouda, Krämer, & Horsels, 2018; Tomesen, Engelen, & Hiddink, 2019; Tomesen, Wouda, Krämer, & Horsels, 2019; Van Til et al., 2018). Average R_{it}-scores varied between 0.35 and 0.47, whereas measurement of accuracy values ranged from 0.86 to 0.97. According to Evers et al. (2010), a reliability coefficient of 0.80 or higher can be considered as "good."

Data Analyses

Hierarchical multiple regression analyses were conducted to examine predictor variables for singing accuracy, melodic achievement, and rhythmic performance abilities (Research Question 1). Models with forced entry of independent variables were constructed to predict their relationship with the dependent variable. All control variables (age, gender, school type, music education, and cognitive levels) and the sum score of the ITAT were entered as a separate block. For multiple regression analyses, linearity was assessed by partial regression plots and a plot of studentized residuals against the predicted values. In all analyses, there was homoscedasticity, as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values. No evidence of multicollinearity, as assessed by tolerance values greater than 0.2, could be found. There were no studentized deleted residuals greater than ±3 *SD*, no leverage values greater than 0.2, and no values for Cook's distance above 1. The assumption of normality was met, as assessed by Q-Q plots. There was independence of residuals as assessed by Durbin-Watson statistics for each model, ranging from 1.684 to 2.005.

Descriptive statistics were used to analyze patterns in vocal performances at the item level (Research Question 2). A Cochran's Q test (Cochran, 1950) was run to determine if the percentages of children transposing a melody was different at each performance trial. Sample size was adequate to use the χ^2-distribution approximation (Tate & Brown, 1970). Kruskall-Wallis H tests were conducted on several VMAT melodic achievement components to determine differences in ability levels by age or grade.

Results

Descriptive statistics of the ITAT, VMAT, and Cito (Dutch school exams) are displayed in online Table S3.

Which Factors (Age, Gender, Cognitive Levels, School Type, Music Perception Abilities, and Participation in Music Education) Predict Vocal Abilities (Singing Accuracy, Melodic Achievement, and Rhythmic Performance)?

Several hierarchical multiple regression analyses were run to determine if the blockwise addition of independent variables improved the prediction of (a) singing accuracy, (b) melodic achievement, and (c) rhythmic performance.

Singing accuracy. The full model (Model 2) was statistically significant, $R^2 = .302$, $F(11, 221) = 8.711$, $p < .0005$, adjusted $R^2 = .268$, $f^2 = 0.433$ (see Table 1). Music perception abilities were significantly and positively associated with singing accuracy. Gender was significant, indicating being a female participant was associated with higher singing accuracy. Furthermore, math was significantly and negatively associated with singing accuracy. Taking both instrumental and singing lessons disappeared as a significant predictor after the ITAT sum score was added in Model 2.

Melodic achievement. The full model (Model 2) was statistically significant, $R^2 = .319$, $F(11, 221) = 9.414$, $p < .0005$, adjusted $R^2 = .285$, $f^2 = 0.468$ (see Table 1). Music perception abilities were significantly and positively associated with melodic achievement, whereas math was significantly negatively associated. Taking both instrumental and singing lessons disappeared as a significant predictor after the ITAT sum score was added in Model 2. Age became a significant predictor only after adding the ITAT sum score in Model 2.

Rhythmic performance. The full model (Model 2) was statistically significant, $R^2 = .429$, $F(18, 214) = 15.100$, $p < .0005$, adjusted $R^2 = .401$, $f^2 = 0.751$ (see Table 1). Music perception abilities and taking instrumental music lessons were significantly and positively associated with rhythmic performance. Age was significant, indicating being older is associated with higher rhythmic performance. School type and taking both instrumental and singing lessons disappeared as significant predictors after the ITAT sums core was added in Model 2.

Do Children of Different Ages Perform Differently With Regard to Subcategories of Vocal Abilities, Such as Transposing Melodies, Reproducing Rhythmic Patterns, Beat-Keeping, Completing Melodic Patterns, and Audiating Tonal Distances?

Significant results are presented by VMAT task; reported percentages refer to the total number of participants ($N = 231$) unless otherwise stated.

Transpose a melody. Participants had to vocally transpose the beginning four notes of a well-known melody for five performance trials on a given different starting tone. This task was done correctly once by 28.8% of the participants, twice by 9.9%, three times by 4.3%, four times by 3.4%, and five times by 0.9%. Cochran's Q test was run to determine if the percentages of children transposing the melody correctly was significantly different between the performance trials. In the first trial, 27.5% ($n = 64$) succeeded; in the second trial, 35.6% ($n = 83$) succeeded; in the third trial, 28.8% ($n = 67$) succeeded; in the fourth trial, 18.0% ($n = 42$) succeeded; and in the fifth trial, 12.0% ($n = 28$) succeeded. The percentage of children vocally transposing the melody correctly was statistically significantly different for different performance trials, $\chi^2(4) = 73.004$, $p < .0005$. Paired comparisons revealed statistically significant differences

Table 1. Predictors of VMAT Singing Accuracy, Melodic Achievement, and Rhythmic Performance in Hierarchical Multiple Regression Models.

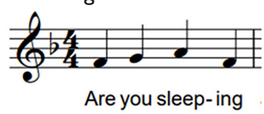
	VMAT Singing Accuracy		VMAT Melodic Achievement		VMAT Rhythmic Performance	
Predictors by Model	Adjusted R^2	β	Adjusted R^2	β	Adjusted R^2	β
Model 1	.161		.136		.344	
Age		0.089		0.027		0.419***
Gender[a]		0.233***		0.122		0.019
School type[b]		–0.107		–0.231***		–0.152**
Instrumental education[c]		0.115		0.087		0.156**
Vocal education[c]		0.000		0.007		0.072
Instrumental & vocal education[c]		0.169*		0.222***		0.243***
Reading accuracy (Cito)		0.031		0.095		0.026
Reading comprehension (Cito)		0.135		0.134		0.046
Math (Cito)		–0.193**		–0.161*		0.038
Grammar (Cito)		0.097		0.011		0.108
Model 2	.268		.285		.401	
Age		–0.130		–0.231**		0.259***
Gender[a]		0.186**		0.066		–0.016
School type[b]		–0.028		–0.137*		–0.094
Instrumental education[c]		0.061		0.022		0.119*
Vocal education[c]		–0.037		–0.032		0.084
Instrumental & vocal education[c]		0.097		0.153		0.065
Reading accuracy (Cito)		0.010		0.070		0.010
Reading comprehension (Cito)		0.077		0.065		0.004
Math (Cito)		–0.198*		–0.167*		0.034
Grammar (Cito)		0.064		–0.028		0.083
ITAT sum score		0.437***		0.516***		0.321***

Note. VMAT = Vocal Musical Ability Test; Cito = Dutch school exams; ITAT = Implicit Tonal Ability Test.
[a]Male = 0; female = 1.
[b]Waldorf school = 0; Dutch public school = 1.
[c]Compared to no music education.
$*p < .05$. $**p < .01$. $***p < .001$.

between several performance trials of transposition tasks (see Table 2). Furthermore, distributions of transposing scores were not similar for all age groups, as assessed by visual inspection of a boxplot. However, the mean rank of successful transposing scores was not statistically significantly different between participants with different ages, $\chi^2(5) = 6.838$, $p = .233$.

Reproduce rhythmic patterns. Participants had to reproduce 10 rhythmic patterns by tapping with their hand of preference. Median age scores differed statistically significantly between grades, $\chi^2(5) = 46.444$, $p < .0005$. Post hoc analysis revealed

Table 2. Pairwise Comparisons of Different Performance Trials, Descriptive Statics, and Original Melody Fragment.

Post Hoc Comparisons					Descriptive Statistics		
Pair[a]	Standard Test Statistic	Significance	Adjusted Significance	Melody Trial	Starting Tone	*M*	*SD*
1–2	–2.640	.008	.083	1	$F\sharp_4$	.27	.447
1–3	–0.417	.677	1.000	2	C_4	.36	.480
1–4	3.057	.002	.022	3	$E\flat_4$	.29	.454
1–5	5.002	$< .001$	$< .001$	4	A_4	.18	.385
2–3	2.223	.026	.262	5	$B\flat_4$	.12	.326
2–4	5.697	$< .001$	$< .001$				
2–5	7.642	$< .001$	$< .001$		Original melody fragment		
3–4	3.474	$< .001$	$< .001$				
3–5	5.419	$< .001$	$< .001$				
4–5	1.945	.052	.518				

Note. Significance values have been adjusted by the Bonferroni correction for multiple tests. $N = 233$.
[a]Pair = combination of melody trials of transposition tasks. For example, Pair 2–4 represents a post hoc comparison between the second and the fourth performance trials.

statistically significant differences in median age scores between grade 1 and grades 3 to 6 ($p < .05$) and between grade 2 and grades 4 to 6 ($p < .01$); see median age values and post hoc results in online Table S4.

Synchronizing and beat-keeping. Participants had to recognize and synchronize to the beat of audio samples of several music pieces by tapping with their hand. Distributions were not similar for all groups ($N = 233$), as assessed by visual inspection of a boxplot. Reported values are mean ranks, and these were statistically significantly different between ages, $\chi^2(5) = 35.677$, $p < .0005$. Post hoc analysis revealed statistically significant differences between age 6 (87.95) and ages 11–12 (162.32; $p < .0005$), between age 7 (84.54) and ages 11–12 ($p < .0005$), and between age 8 (113.75) and ages 11–12 ($p = .019$).

Complete melodic patterns. Participants had to sing the following "logical" tone after listening to five melodic patterns. Descriptive statistics of successful pattern completion by age are displayed in online Table S5. Median scores were statistically significantly different between ages, $\chi^2(5) = 15.099$, $p = .010$. Post hoc analysis revealed statistically significant differences in melodic pattern completion between age 6 ($Mdn = 1.00$) and ages 11–12 ($Mdn = 2.00$, $p = .021$) only.

Table 3. Results in Percentages of Tonal Distance Performance Tasks, Specified by Outcome.

Audiation of Tonal Distances on Starting Tone F_4	1 Tone ↑	2 Tones ↑	3 Tones ↑	1 Tone ↓	2 Tones ↓
Singing a unison	22.7	11.2	12.4	18.5	15.0
Singing correct tone (major interval)	24.9	21.0		4.3	2.1
Singing correct tone (minor interval)	1.7	2.6		5.2	12.9
Singing correct tone (perfect interval)			12.4		
Singing correct tone (augmented interval)			3.4		
Singing perfect interval (except unison) in either direction	16.7	24.3	27.5[a]	38.6	37.4
Singing other interval in right direction	24.1	31.0	33.9	19.8	20.2
Singing other interval in wrong direction	9.8	9.9	10.2	13.7	12.4

Note. Percentages are rounded to one decimal. ↑ = higher; ↓ = lower.
[a]Percentage excludes correct outcome for ascending perfect fourth.

Audiate tonal distances. Participants had to audiate and consequently sing tonal distances in ascending and descending direction. A single tone (F_4) was given as a starting point. Relatively few children were able to audiate tonal distances and sing the correct interval based on mental representation. Moreover, descending intervals were more likely to be sung incorrectly than ascending intervals. Approximately 10% of children confused higher and lower, and even more children just repeated the starting tone. On average, 30% of produced intervals were a perfect fourth, fifth, or eighth, despite the instructions. Percentages specified by outcome are displayed in Table 3.

Discussion

This study set out to investigate associations between implicit tonal abilities, cognitive levels, and vocal abilities in elementary school children. The first research question examined which factors (age, gender, cognitive levels, school type, music perception abilities, and participation in music education) predict vocal abilities (singing accuracy, melodic achievement, and rhythmic performance). As our results suggested, abilities can differ greatly between children of the same age, even to the extent that children age 6 can easily outperform children age 12. If we analyze the role of age, we see a remarkable development. The negative β for age in the model for predicting melodic achievement suggested that children sing at their best around the ages of 9 and 10. This age-related decline did not happen in the singing accuracy component, although boys of all ages sang less accurately. It is therefore also unlikely that the older boys' development of vocal cords is associated with the overall decline. Age-related inhibition to sing in public may explain our findings. Evidently, girls are, on average, better at using their singing voice and intonating pitches than boys. From a pedagogical point of view, these findings would argue in favor of age grouping instead of grouping according to vocal skills. For educators, an answer to the question why vocal abilities decline from the age of 11 is relevant. We reason that the ability to sing in tune might mainly depend on the quality of hearing and (gender-related) usable vocal registers rather than being the result of vocal training or cognitive maturation, as was also proposed by Levinowitz et al. (1998).

Furthermore, it is important to note that the vast majority of children were untrained, which may have influenced our results. Therefore, we are cautious not to underestimate the role of vocal training. For rhythmic performance abilities, our findings suggested an effect of formal instrumental training and maturation (growth in both memory and motor skills). Moreover, children with better musical cognition may have better hearing and intonation skills and a wider usable vocal register. Consequently, musically gifted children might engage in musical activities more frequently than their less musically gifted peers. Another notable result to discuss is the negative association of math level on children's vocal abilities. To the best of our knowledge, there is no existing literature consistent with our findings. While gender did show significant differences in the singing accuracy component, this is not the case for melodic achievement and rhythmic performances. Therefore, we conclude that gender does not affect the quality of musical cognition in relation to vocal abilities. Aspects such as memory and motor skills that are important for rhythmic pattern reproduction also appear to be gender-independent.

The second research question examined if children of different ages perform differently with regard to subcategories of vocal abilities, such as transposing melodies, reproducing rhythmic patterns, beat-keeping, completing melodic patterns, and audiating tonal distances. Based on the absent role of age in singing accuracy tasks, we conclude that pitch-matching is not related to maturation, at least not in the age range of 6 to 12. This implies that elementary school children, regardless of their age, varied similarly in their ability sing in tune. In line with this finding, age did not predict achievement on the repeating scales and repeating melodic patterns task. However, three other melodic achievement tasks did differ by age or were hard to perform for any age at all. What makes completing a melodic pattern, audiating tonal distances, and transposing a melody different from repeating scales and melodic patterns? First of all, children could not simply repeat what they had previously heard, as was the case in repeating scales and melodic patterns. In fact, children had to use their musical cognition to complete these musical fragments. We discuss these tasks in the following.

Melodic pattern completion was done significantly better by children ages 11 and 12 compared to age 6. These findings are in line with the general idea about enculturation as a process that takes place in early and middle childhood: Older children gain a better understanding of the structure of their culture's music due to their improved cognitive abilities (Corrigall & Schellenberg, 2015). The low success rates of audiating tonal distances and transposing a melody indicated that sizes of minor and major seconds as the equivalent of a musical distance measure of one tone are far from common in children's mental representations. Familiarity with or preference for singing perfect intervals might be the reason that children sang so many perfect fourths and fifths.

Age did also play a role in tasks of the rhythmic performance component. Especially for the younger children until the age of 8 (grade 3), less developed motor control, coordination, and memory skills might explain the lower success rates for repeating rhythmic patterns, as is in line with previous research (see also Drake et al., 2000). Recognizing beat, synchronizing to it, and beat-keeping improved until the age of 9. Comparable to some extent are findings from Slater et al. (2013) that beat-keeping is a matter of experience.

Conclusions, Implications, Limitations, and Future Research

This study has provided new insights into vocal abilities of elementary school children in relation to musical perception abilities and music education, cognitive levels, gender, and age. In summary, music perception abilities were the most important for predicting both vocal and rhythmic abilities. Singing accuracy was affected by gender, while melodic achievement and rhythmic performance were affected by age. Furthermore, formal musical training only contributed to rhythmic performance skills. Melodic achievement components did not develop uniformly. Audiating tonal distances and transposing melodies was difficult for children of all ages. Rhythmic performance skills increased up to the age of 8 years, while beat-keeping increased up to 9 years. Successful completion of melodic patterns gradually increased in the age range studied but only significantly differed between the youngest and oldest children. There was no significant relationship between vocal abilities and academic cognitive levels, except for a negative association between math and singing accuracy and melodic achievement. Our findings emphasize the importance of musical perception training of elementary school children as well as a curriculum in which time is devoted to musical activities.

In the lower grades of elementary school, it is not necessary to avoid songs which contain large intervals. Therefore, we advise teachers to let all children regardless of their age sing songs and tonal patterns that contain a variety of intervals (including those larger than an octave). After all, we did not find any significant differences by children's age for repeating small and large intervals. Incidentally, this advice is not a call to offer extreme exercises that could damage children's voices due to strenuous demands. Furthermore, the quality of singing education in elementary schools can be boosted by offering various singing tasks in which repetition and prediction of patterns, scales, intervals, and intonation are practiced as well as transposing melodies and repeating rhythmic patterns. Simply memorizing songs or singing along to a tune may be the most popular activity in elementary music education. However, it would be a sign of a poor curriculum or a lack of training if teachers do not engage in other activities. Using various singing tasks that include audiation skills are also recommended by closely related research (Nichols, 2017; Reifinger, 2020).

Due to the design of our research, we cannot make any statements about causal relations between perception abilities and vocal abilities. Moreover, independent validation of our measurement instruments by other researchers in other countries could strengthen our results. In addition, differences in music curricula and culture worldwide caution us not to generalize beyond the local region or national borders. More research is needed to better understand how cognition and maturation, perception of musical features, and vocal abilities develop during childhood. Our findings can help music educators to design music methods for elementary education that match the characteristics of musical development of children. With regard to vocal training, elementary school education curricula should pay attention to singing accuracy, melodic achievement, and rhythmic performance by trying to expand children's usable vocal register, enhance their musical cognition, and improve their psychomotor skills. This can be done by offering a total musical package in which perception of all musical features and production activities, such as singing and percussion, are offered. Music seems to be one of the few subjects in elementary school where academic cognitive development does not play a major role. Considering this rare situation, musical activities should be enjoyable for all students. However, the large differences in levels must be taken into account. Too much focus on results will deter children. After all, crows may never sing like nightingales.

Acknowledgments

The authors wish to thank all 233 children, their parents, and their teachers who took part in or cooperated with this study. Furthermore, we would like to thank the coder (J. H. L. Dols) and the English language editor (P. J. Morales).

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: This work is part of the research program Doctoral Grant for Teachers with Project No. 023.012.020, which is financed by the Dutch Research Council (NWO).

Ethics Approval Statement

Approval was obtained from the Research Ethics Committee (cETO) of the Open University of the Netherlands (approval ID U2018/03755/HVM).

ORCID iD

Zyxcban G.Wolfs https://orcid.org/0000-0003-0353-2881

Supplemental Material

Appendix SA, Tables S1 through S5, and Figure S1 are available in the online version of the article at https://doi.org/10.1177/00224294221134571.

Data Availability Statement

The processed data that support the findings of this study are openly available in KNAW EASY at https://doi.org/10.17026/dans-xrh-ayu6. The raw data of this study are not publicly available due to privacy restrictions.

References

Bergeson, T. R., & Trehub, S. E. (2006). Infants' perception of rhythmic patterns. *Music Perception*, *23*(4), 345–360. https://doi.org/10.1525/mp.2006.23.4.345

Berkowska, M., & Dalla Bella, S. (2013). Uncovering phenotypes of poor-pitch singing: the Sung Performance Battery (SPB). *Frontiers in Psychology*, *4*, Article 714. https://doi.org/10.3389/fpsyg.2013.00714

Central Institute of Test Development. (2021). *We serve primary education*. Cito. https://www.cito.com/we-serve/primary-education

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, *37*(3/4), 256–266. https://doi.org/10.2307/2332378

Cohen, A. J. (2015). The AIRS Test Battery of Singing Skills: Rationale, item types, and lifespan scope. *Musicae Scientiae*, *19*(3), 238–264. https://doi.org/10.1177/1029864915599599

Corrigall, K. A., & Schellenberg, G. E. (2015). Music cognition in childhood. In G. E. McPherson (Ed.), *The child as musician: A handbook of musical development* (2nd ed., pp. 81–101). Oxford University Press.

Corrigall, K. A., & Trainor, L. J. (2014). Enculturation to musical pitch structure in young children: Evidence from behavioral and electrophysiological measures. *Developmental Science*, *17*(1), 142–158. https://doi.org/10.1111/desc.12100

Cuadrado, A., & Rusinek, G. (2016). Singing and vocal instruction in primary schools: An analysis from six case studies in Spain. *British Journal of Music Education*, *33*(1), 101–115. https://doi.org/10.1017/S0265051715000273

Demorest, S. M., Nichols, B. E., & Pfordresher, P. Q. (2018). The effect of focused instruction on young children's singing accuracy. *Psychology of Music*, *46*(4), 488–499. https://doi.org/10.1177/0305735617713120

Demorest, S. M., & Pfordresher, P. Q. (2015). Singing accuracy development from K-adults: A comparative study. *Music Perception: An Interdisciplinary Journal*, *32*(3), 293–302. https://doi.org/10.1525/mp.2015.32.3.293

Drake, C., & Gérard, C. (1989). A psychological pulse train: How young children use this cognitive framework to structure simple rhythms. *Psychological Research*, *51*(1), 16–22. https://doi.org/10.1007/BF00309271

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*(3), 251–288. https://doi.org/10.1016/s0010-0277(00)00106-2

Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2010). *COTAN Beoordelingssysteem voor de kwaliteit van tests* [COTAN Quality rating system of tests]. NIP/COTAN.

Gordon, E. E. (2012). *Learning sequences in music*. GIA.

Gould, A. O. (1969). Developing specialized programs for singing in the elementary school. *Research in Music Education*, *17*, 9–22. http://www.jstor.org/stable/40317056

Hedden, D. (2012). An overview of existing research about children's singing and the implications for teaching children to sing. *Update: Applications of Research in Music Education*, *30*(2), 52–62.

Hop, M., & Engelen, R. (2017). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 6* [Scientific accountability Arithmetic-Mathematics 3.0 for grade 4]. Cito.

Hop, M., Janssen, J., & Engelen, R. (2016). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 5* [Scientific accountability Arithmetic-Mathematics 3.0 for grade 3]. Cito.

Hop, M., Janssen, J., & Engelen, R. (2017). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 7* [Scientific accountability Arithmetic-Mathematics 3.0 for grade 5]. Cito.

Hop, M., Scheltens, F., & Engelen, R. (2019). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 8* [Scientific accountability Arithmetic-Mathematics 3.0 for grade 6]. Cito.

Janssen, J., Hop, M., & Wouda, J. (2015). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 4* [Scientific accountability Arithmetic-Mathematics 3.0 for grade 2]. Cito.

Janssen, J., Hop, M., Wouda, J., & Hollenberg, J. (2015). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 3* [Scientific accountability Arithmetic-Mathematics 3.0 for grade 1]. Cito.

Jolink, A., Tomesen, M., Hilte, M., Weekers, A., & Engelen, R. (2015a). *Wetenschappelijke verantwoording Begrijpend lezen 3.0 voor groep 3* [Scientific accountability Reading Comprehension 3.0 for grade 1]. Cito.

Jolink, A., Tomesen, M., Hilte, M., Weekers, A., & Engelen, R. (2015b). *Wetenschappelijke verantwoording Begrijpend lezen 3.0 voor groep 4* [Scientific accountability Reading Comprehension 3.0 for grade 2]. Cito.

Kisilevsky, B. S., Hains, S. M. J., Jacquet, A.-Y., Granier-Deferre, C., & Lecanuet, J. P. (2004). Maturation of fetal responses to music. *Developmental Science*, *7*(5), 550–559. https://doi.org/10.1111/j.1467-7687.2004.00379.x

Levinowitz, L. M., Barnes, P., Guerrini, S., Clement, M., D'April, P., & Morey, M. J. (1998). Measuring singing voice development in the elementary general music classroom. *Journal of Research in Music Education*, *46*(1), 35–47. https://doi.org/10.1525/mp.2015.32.3.283

Mang, E. (2006). The effects of age, gender and language on children's singing competency. *British Journal of Music Education*, *23*(2), 161–174. https://doi.org/10.1017/S0265051706006905

Matsunaga, R., Hartono, P., & Abe, J. (2015). The acquisition process of musical tonal schema: Implications from connectionist modeling. *Frontiers in Psychology*, *6*, Article 1348. https://doi.org/10.3389/fpsyg.2015.01348

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348–367. https://doi.org/10.1037/0096-3445.135.3.348

Morrison, S. J., Demorest, S. M., & Stambaugh, L. A. (2008). Enculturation effects in music cognition. The role of age and music complexity. *Journal of Research in Music Education*, *56*(2), 118–129. https://doi.org/10.1177/0022429408322854

Nichols, B. E. (2016a). Critical variables in singing accuracy test construction: A review of literature. *Update: Applications of Research in Music Education*, *35*(1), 39–46. https://doi.org/10.1177/8755123315576764

Nichols, B. E. (2016b). Task-based variability in children's singing accuracy. *Journal of Research in Music Education*, *64*(3), 309–321. https://doi.org/10.1177/0022429416666054

Nichols, B. E. (2017). Constructing singing assessments for the music classroom. *General Music Today*, *30*(3), 13–17. https://doi.org/10.1177/1048371317690864

Nichols, B. E. (2020). Effect of vocal versus piano doubling on children's singing accuracy. *Psychology of Music*, *49*(5), 1415–1423. https://doi.org/10.1177/0305735620936757

Pearce, M., & Rohrmeier, M. (2012). Music cognition and the cognitive sciences. *Topics in Cognitive Science*, *4*(4), 468–484. https://doi.org/10.1111/j.1756-8765.2012.01226.x

Peretz, I., Gosselin, N., Nan, Y., Caron-Caplette, E., Trehub, S. E., & Béland, R. (2013). A novel tool for evaluating children's musical abilities across age and culture. *Frontiers in Systems Neuroscience*, *7*, Article 30. https://doi.org/10.3389/fnsys.2013.00030

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of "tone deafness." *Music Perception*, *25*(2), 95–115. https://doi.org/10.1525/mp.2007.25.2.95

Pfordresher, P. Q., & Demorest, S. M. (2021). The prevalence and correlates of accurate singing. *Journal of Research in Music Education*, *69*(1), 5–23. https://doi.org/10.1177/0022429420951630

Philips, K. H., & Doneski, S. M. (2012). Research on elementary and secondary school singing. In R. Colwell & P. Webster (Eds.), *MENC handbook of research on music learning: Vol.2 applications* (pp. 176–232). Oxford University Press.

Reifinger, J. L., Jr. (2020). Teaching pitch notation-reading skills. *General Music Today*, *33*(3), 21–28. https://doi.org/10.1177/1048371319891419

Roberts, E., & Davies, A. (1975). Poor pitch singing: Response of montone singers to a program of remedial training. *Journal of Research in Music Education*, *23*(4), 227–239. https://doi.org/10.2307/3344852

Rutkowski, J. (2015). The relationship between children's use of singing voice and singing accuracy. *Music Perception: An Interdisciplinary Journal*, *32*(3), 283–292. https://doi.org/10.1525/mp.2015.32.3.283

Rutkowski, J. (2019). Assessing singing voice development. In T. S. Brophy (Ed.), *The Oxford handbook of assessment policy and practice in music education* (*Vol. 2*, pp. 629–652). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190248130.013.64

Salvador, K. (2010). How can elementary teachers measure singing voice achievement? A critical review of assessments, 1994–2009. *Update: Applications of Research in Music Education*, *29*(1), 40–47. https://doi.org/10.1177/8755123310378454

Slater, J., Tierney, A., & Kraus, N. (2013). At-risk elementary school children with one year of classroom music instruction are better at keeping a beat. *PLOS ONE*, *8*(10), Article e77250. https://doi.org/10.1371/journal.pone.0077250

Speer, J. R., & Meeks, P. U. (1985). School children's perception of pitch in music. *Psychomusicology*, *5*(1), 49–56. https://doi.org/10.1037/h0094200

Szabo, M. (2001). Children's inaccurate singing: Selected contributing factors. *General Music Today*, *14*(3), 4–10. https://doi.org/10.1177/104837130101400303

Tate, M. W., & Brown, S. M. (1970). Note on the Cochran Q test. *Journal of the American Statistical Association*, *65*(329), 155–160. https://doi.org/10.2307/2283582

Tomesen, M., Engelen, R., & Hiddink, L. (2018). *Wetenschappelijke verantwoording Begrijpend lezen 3.0 voor groep 7* [Scientific accountability Reading Comprehension 3.0 for grade 5]. Cito.

Tomesen, M., Engelen, R., & Hiddink, L. (2019). *Wetenschappelijke verantwoording Begrijpend lezen 3.0 voor groep 8* [Scientific accountability Reading Comprehension 3.0 for grade 6]. Cito.

Tomesen, M., Weekers, A., Hiddink, L., & Jolink, A. (2017). *Wetenschappelijke verantwoording Begrijpend lezen 3.0 voor groep 6* [Scientific accountability Reading Comprehension 3.0 for grade 4]. Cito.

Tomesen, M., Weekers, A., Hilte, M., Jolink, A., & Engelen, R. (2016). *Wetenschappelijke verantwoording Begrijpend lezen 3.0 voor groep 5* [Scientific accountability Reading Comprehension 3.0 for grade 3]. Cito.

Tomesen, M., Wouda, J., & Horsels, L. (2016). *Wetenschappelijke verantwoording van de LVS-toetsen Spelling 3.0 voor groep 5* [Scientific accountability Grammar 3.0 for grade 3]. Cito.

Tomesen, M., Wouda, J., & Horsels, L. (2017). *Wetenschappelijke verantwoording van de LVS-toetsen Spelling 3.0 voor groep 6* [Scientific accountability Grammar 3.0 for grade 4]. Cito.

Tomesen, M., Wouda, J., Krämer, I., & Horsels, L. (2018). *Wetenschappelijke verantwoording van de LVS-toetsen Spelling 3.0 voor groep 7* [Scientific accountability Grammar 3.0 for grade 5]. Cito.

Tomesen, M., Wouda, J., Krämer, I., & Horsels, L. (2019). *Wetenschappelijke verantwoording van de LVS-toetsen Spelling 3.0 voor groep 8* [Scientific accountability Grammar 3.0 for grade 6]. Cito.

Tomesen, M., Wouda, J., Mols, A., & Horsels, L. (2015a). *Wetenschappelijke verantwoording van de LVS-toetsen Spelling 3.0 voor groep 3* [Scientific accountability Grammar 3.0 for grade 1]. Cito.

Tomesen, M., Wouda, J., Mols, A., & Horsels, L. (2015b). *Wetenschappelijke verantwoording van de LVS-toetsen Spelling 3.0 voor groep 4* [Scientific accountability Grammar 3.0 for grade 2]. Cito.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 394–402. https://doi.org/10.1037/0096-1523.18.2.394

Ullén, F., Mosing, M. A., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. *Personality and Individual Differences*, *63*, 87–93. https://doi.org/10.1016/j.paid.2014.01.057

Van Til, A., Kamphuis, F., Keuning, J., Gijsel, M., Vloedgraven, J., & De Wijs, A. (2018). *Wetenschappelijke verantwoording LVS-toetsen DMT* [Scientific accountability Reading Accuracy]. Cito.

Welch, G. F. (2015). Singing and vocal development. In G. McPherson (Ed.), *The child as musician: A handbook of musical development* (2nd ed., pp. 441–461). Oxford University Press.

Welch, G. F., Sunders, J., Papageorgi, I., & Himonides, E. (2012). Sex, gender and singing development: Making a positive difference to boys' singing through a national programme in England. In S. Harrison, G. Welch, & A. Adler (Eds.), *Perspectives on males singing. Landscapes: The arts, aesthetics, and education* (*Vol. 10*, pp. 27–43). Springer. https://doi.org/10.1007/978-94-007-2660-4_3

Author Biographies

Zyxcban G. Wolfs is an elementary school teacher, musician, and researcher at the Open University of the Netherlands. His research interests include music perception abilities and music performance skills.

Henny P. A. (Els) Boshuizen is professor emerita of educational sciences at the Open University of the Netherlands. Her research projects and publications span the development of professional expertise in several domains, the associated learning processes and required cognitive changes, and educational strategies to support learners on their way to expertise.

Saskia Brand-Gruwel is a professor in learning sciences at the Open University of the Netherlands. Her research interests include instructional design, information problem solving, and self-regulated learning.

Submitted January 28 2022; accepted August 8, 2022.