Article

National Association for Music Education

Update
1–8
© 2015 National Association for Music Education
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/8755123315576764
update.sagepub.com
SAGE

Critical Variables in Singing Accuracy Test Construction: A Review of Literature

Bryan E. Nichols[1]

Abstract

The purpose of this review of literature was to identify research findings for designing assessments in singing accuracy. The aim was to specify the test construction variables that directly affect test performance to guide future design in singing accuracy assessment for research and classroom uses. Three pitch-matching tasks—single pitch, interval pitch, and patterns—plus song-singing have each been shown to be discriminators of singing accuracy. Teachers and researchers should consider incorporating more than one task type when assessing children, since performance on one song or pitch-matching task cannot easily be generalized to others. Item-level and task-level difficulty must be taken into consideration when designing singing tests. In addition to task type, the variables of singing voice development, contextual presentation, model characteristics, range, and text must be a part of a priori decision making.

Keywords

assessment, in-tune singing, music test, pitch-matching, singing accuracy

In many, if not all general music classrooms, singing is a primary vehicle for instruction. Even at the secondary level, when students participate in performance-based music classes such as band, choir, and orchestra, a student's ability to sing accurately may indicate overall musicianship or may demonstrate music difficulties that the student is having. It is often the goal of the music teacher to ensure that students can sing accurately for success in the classroom and for future music success. The development of children's singing accuracy is a responsibility of music teachers, and it is important for teachers to have comprehensive measures of singing accuracy.

Researchers in music education have created unique measures of singing (see Salvador, 2010, for a review) and have tested many different elementary grade samples to describe properties like singing range, accuracy, and other features (Hedden, 2012). Children have demonstrated widely varying competencies, from only 9% to 38% of children deemed inaccurate (Goetze, Cooper, & Brown, 1990) to 75% to 90% defined as nonsingers (Levinowitz et al., 1998). However, these figures reflect samples of students of different age ranges or use test items that vary in difficulty. Additionally, different models of evaluation and scoring are applied to these assessments, which in turn further obscure the true percentage of the student population with undeveloped or underdeveloped singing skills. In other words, the prevalence of accurate singing depends on the difficulty of the test items, the type of measure used for evaluation, and the definition of accurate singing.

There are two main areas of study that are relevant to singing accuracy, which is also referred to as vocal pitch accuracy. First are presentation variables, which are the effects of the type of singing task and how singing test items are presented to students. Second are response variables, which are the specific conditions under which students are asked to respond. This article will review the music education literature on presentation and response variables, which generally reports on school-aged children for participation, rather than adults. References will be made to studies outside of music education when the findings are relevant to children's singing.

Singing Tasks Used in Assessment

Students are asked to sing in music classroom and rehearsal situations in a variety of ways. Specific pitch sequences (referred to as test items when used in an assessment) can be grouped into types of tasks in two response modes. First, students can be asked to echo given pitches and second, students can be asked to recall

[1]The University of Akron, OH, USA

Corresponding Author:
Bryan E. Nichols, The University of Akron, Guzzetta Hall 254, Akron, OH 44325-1002, USA.
Email: bnichols@uakron.edu

Downloaded from upd.sagepub.com by guest on April 17, 2015

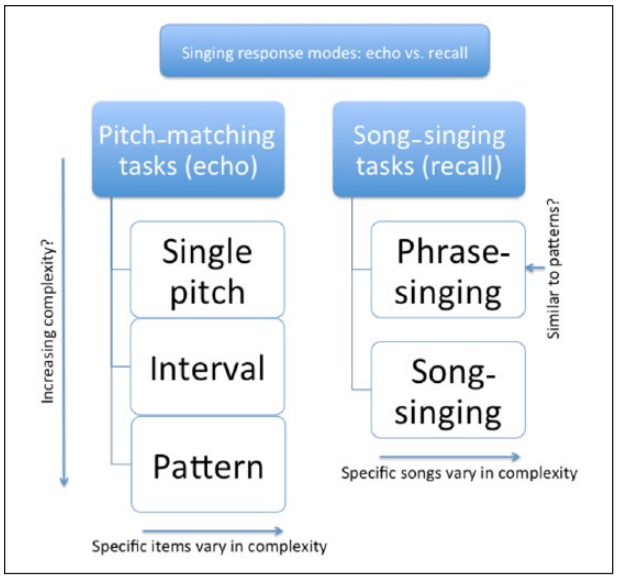

Figure 1. Singing response modes.
Note. This figure illustrates the different types and relationships among singing tasks.

melodic material from memory rather than echo (see Figure 1). For assessment, these response modes are represented by pitch-matching tasks and song-singing tasks, respectively. Within pitch-matching, students can be given a single pitch, an interval, or patterns of varying lengths. For song singing, students can be asked to recall songs or fragments of songs, like phrases.

Evidence has supported the use of tonal patterns and familiar songs as assessment tasks (Phillips & Doneski, 2011). Pitch-matching and song-singing have been used extensively, and they have sometimes been combined in assessments. However, when this has occurred, few comparisons among tasks have been made. The use of a single pitch for the pitch-matching task is briefest and has been used for singing accuracy assessment (Demorest, 2001; Demorest & Clements, 2007; Flowers & Dunne-Sousa, 1990; Geringer, 1983; Jones, 1971; Murry & Zwirner, 1991; Nichols, 2013; Porter, 1977; Roberts & Davies, 1975). Geringer (1983), as well as Demorest and Clements (2007), used single pitch tasks presented in a specific tonal context, which has not been explored in other studies. The choice of intervals as test items in singing assessments has been popular (Green, 1990; Moore, Fyk, Frega, & Brotons, 1995; Price, Yarbrough, Jones, & Moore, 1994; Yarbrough, Bowers, & Benson, 1992; Yarbrough, Green, Benson, & Bowers, 1991) and has been used as a part of larger assessments incorporating additional tasks (Flowers & Dunne-Sousa, 1990; Joyner, 1969; Klemish, 1974; Nichols, 2013; Roberts & Davies, 1975). Melodic patterns of three or more notes have also been used exclusively to inform singing accuracy performance (Cooper, 1995; Phillips & Aitchison, 1997, 1999; Phillips, Aitchison, & Nompula, 2002; Rutkowski, 1990, 1996) or as a part of a larger assessment (Flowers & Dunne-Sousa, 1990; Guerrini, 2006; Nichols, 2013; Welch, Sergeant, & White, 1997). Researchers have designed one-measure song fragments to investigate singing voice development, a different construct from singing accuracy that primarily includes vocal features such as voice range (Levinowitz et al., 1998; Rutkowski, 1990, 1996; Rutkowski & Miller, 2003).

Play songs and folk songs have been the test stimuli for many children's studies (Apfelstadt, 1984; Brophy, 1997; Flowers & Dunne-Sousa, 1990; Gault, 2002; Goetze, 1985; Green, 1994; Guerrini, 2006; Guilbault, 2004; Joyner, 1969; Muse, 1993; Nichols, 2013; Roberts & Davies, 1975; Smith, 1973; Welch et al., 1997; Western, 2002; Wurgler, 1990; Young, 1971). Studies of singing among adults have incorporated similar songs, such as *Twinkle, Twinkle Little Star* (Pfordresher, Brown, Meier, Belyk, & Liotti, 2010), *Row, Row, Row Your Boat* (McCoy, 1997; Pfordresher et al., 2010), *Jingle Bells* (Berkowska & Dalla Bella, 2009; Pfordresher et al., 2010), *Brother John* (Berkowska & Dalla Bella, 2009), *Gens du Pays*, a Quebecois birthday song (Dalla Bella, Giguère, & Peretz, 2007), *Happy Birthday* (Pfordresher et al., 2010), a Polish happy birthday song (Dalla Bella & Berkowska, 2009), *America* (Phillips & Vispoel, 1990; Watts, Moore, & McCaghren, 2005), and *O Music* (Phillips & Vispoel, 1990).

Comparisons of pitch-matching and song-singing tasks are few in number, but they do exist. Five-year-olds have been shown to perform better on single pitch-matching, glides, and patterns than in song-singing (Welch, Sergeant, & White, 1995). These young singers were evaluated on a 7-point scale that was applied in the same way to each of the tasks. Students performed single pitches better than patterns. In a similar study including older students, pitch patterns were again demonstrated to be easier than songs for 5-, 6-, and 7-year-olds, and the authors suggested that accuracy is task-specific (Welch et al., 1997).

Performance on pitch-matching and song-singing was found to be moderately correlated in a sample of fourth graders (Nichols, 2013), and highly correlated in a sample of adults (Pfordresher et al., 2010). Pfordresher and Brown (2007) suggested that single pitch, interval, and pattern tasks vary in complexity because of the number of unique pitches they contain. Furthermore, the authors suggested these types of tasks vary in terms of memory demand because of the total number of notes they contain. Therefore, they sought to equalize demands on memory by presenting the single pitch task in four (identical) notes, the interval pitches in four notes (two for the first pitch and two for the second pitch), and the pattern in

Downloaded from upd.sagepub.com by guest on April 17, 2015

four unique pitches. For those participants deemed good singers, performance improved across these three tasks, respectively. For poor singers, however, performance decreased across these tasks, suggesting that good singers may benefit from the added contextual information, whereas poor singers become more and more overwhelmed with additional complexity. In contrast, Wise and Sloboda (2008) found performance of tone-deaf and not tone-deaf adults to differ significantly from each other, but the performance of both groups across the three tasks (single pitches, intervals, patterns) was similar in that all adults' performances decreased across tasks of additional complexity. However, they did not control for memory by presenting the same number of notes in each pitch-matching task type, which may or may not actually affect accuracy.

Familiarity with specific patterns or a song may not increase the likelihood for improved performance. Guerrini (2006) had fourth- and fifth-grade students sing patterns followed by the songs *America* (familiar) and *Path to the Moon* (unfamiliar). Student responses were significantly higher in the pattern test than in the song-singing tasks, and students performed the familiar and unfamiliar songs with similar accuracy. Guerrini (2006) interpreted these results to mean that students can accurately sing patterns before whole songs, and he makes three conclusions related to singing development:

1. Many students learn to sing accurately by expanding their ranges first.
2. After the range has been expanded, many students become more vocally accurate by singing short patterns, but may not be able to apply this skill to entire songs.
3. Once students are able to sing one song accurately, using notes above the lift, they appear to be able to sing other songs accurately. (p. 167)

Van Zee (1984) also had students sing patterns and a song, *Hot Cross Buns*, in an effort to explore remediation techniques for first graders. After an analysis and description of individual students, she concluded that the ability to "match tonal patterns does not always carry over to singing the songs from which they were taken. A tonal memory needs to be developed for longer phrases" (p. 70). Van Zee suggested that tonal memory plays a role but, more important, that pitch-matching exercises like patterns may not always correlate to functional singing test items like song-singing.

Taking a different approach, Apfelstadt (1984) tested the effect of melodic perception instruction among kindergarteners. Her assessment included the song *Jingle Bells*, in addition to patterns. Pitch-matching performance was significantly improved after a period of instruction, but song-singing was not. The author did not examine how the nature of pitch-matching and song-singing differs; however, these results, combined with others described in this review, lead to the conclusion that the quality of performance on different pitch-matching tasks may differ and may also vary by age. As is suggested in the conclusion, a combination of these tasks may be necessary to test the skill of singing accuracy.

Authors of studies reviewed in this section have suggested that (a) singing accuracy is task-based and (b) future research is warranted because there are important classroom and research implications for individual students across the spectrum of ability/aptitude. Singing accuracy assessment depends on a thorough understanding of how performance differs based on the task given and under which conditions students respond most accurately. Once assessment is better understood, remediation for inaccurate singers can be refined based on the transfer of skills from easy tasks to more difficult tasks.

Item Difficulty

This section is used to present studies that explicitly attempted to define the comparative difficulty of pitch sequences, including discrimination power where available. For specific intervals, there is a perception among teachers that thirds, especially descending thirds, are easiest to sing. However, four-note patterns and half steps have been shown to be no more difficult to sing than whole steps or thirds (Sinor, 1984), although descending third intervals were indexed as easy in many of the patterns. Young (1971) presented intervals in minor and major designations to kindergarten and first graders. He reported a range of intervals from easy to difficult: descending minor third, ascending and descending fifth, ascending major third, and descending minor sixth. In his pilot study using second and third graders, Jones (1971) also listed intervals and patterns in order of difficulty: descending minor third, ascending perfect fourth, *mi-re-do*, ascending second, and *do-re-mi*.

Van Zee (1984) studied students in the first 10 grades and found the ascending fourth, *sol-do*, to be the easiest interval for participants, which corroborates data from Petzold (1963) and Boardman (1964). Van Zee questioned how much pitch-matching tasks correlate with song-singing tasks. She found little or no correlation and suggested they may not measure the same construct (or at least not exactly the same combination of component skills). Perhaps the isolated, even sterilized, task of pitch matching is a skill that can be easily taught, facilitated, and reinforced. The task may be a minimal challenge for short-term memory, and since intervals and patterns are

Downloaded from upd.sagepub.com by guest on April 17, 2015

brief, they may be easier and less intimidating to sing than whole songs.

Tonality was compared in a study of song-singing tasks, and children in Grades 1 through 6 performed the major key song, *Row, Row, Row Your Boat* more accurately than the minor key song, *In the Sea* (Levinowitz et al., 1998). The authors of that study, however, suggested the two songs were not equivalent, thus negating the conclusion that singing in a major key is easier. Earlier, it had been demonstrated twice that singing a minor song was easier than singing a major song (DeYarman, 1972; Dittemore, 1969, as cited in Levinowitz et al., 1998). Thus, there are conflicting results as to whether children sing minor or major songs more accurately.

In a study exploring the effect of tonality on singing accuracy, Wolf (2005) administered the Tonal Performance Pattern Test to kindergarten through second graders in order to determine the difficulty of 40 two- and three-note patterns split between parallel forms of a major key and a minor key section of the test. Seven patterns were easy, 7 were difficult, and the remaining 26 patterns were deemed moderate. Three factors did not seem to affect difficulty:

1. Modality (major or minor)
2. Contour (ascending or descending)
3. Harmonic function (the patterns were classified as Tonic or Dominant in function)

Three factors did seem to determine difficulty in performance:

1. Interval (all easy patterns were thirds)
2. Length (all easy patterns were two notes)
3. Range (all easy patterns were low in range)

The second finding reinforces those from other studies (e.g., Nichols, 2013) that as tasks increase in number of pitches, they may increase in complexity and also difficulty.

A study of 480, 6- through 9-year-olds from four countries was conducted using 16 intervals (Moore et al., 1995). In order to limit the test to 10 minutes, the intervals consisted of a unison, second, third, fourth, fifth, and octave, but not a sixth or seventh. Overall, 60% of the intervals were sung correctly by the participants. Results indicated that the unison, descending minor third, and perfect fourth were easier than the minor second and octave. Specifically, the order of difficulty (pass rate) was unison (80%), descending minor third (73%), descending perfect fourth (68%), ascending major second (66%), descending major third (65%), descending major second (62%), ascending minor third (59%), ascending major third (58%), descending and ascending perfect fifth (both 54%), ascending perfect fourth (51%), descending half step (46%), ascending octave (45%), ascending half step (38%), and descending octave (36%). There were no significant differences between ascending and descending intervals.

Task Discrimination

Roberts and Davies (1975) tested all three pitch-matching task types (single pitch, intervals, patterns), plus song-singing, to study the effects of remediation instruction in a sample of 90, 6- through 8-year-old students identified by the classroom teacher as monotone singers. The sample was randomly assigned to a control group, a traditional group, or a remedial group, with 30 participants in each. An additional group of 30 normal singers was added for comparison. For testing on the song task, children were allowed to select any song or nursery rhyme to sing from memory. After 8 weeks of meeting twice a week for singing improvement sessions, both the normal and monotone singers improved on all measures of singing production. The remedial group indicated better improvement of single pitch and interval production. Though their vocal range improved overall, their song singing did not. Other samples of typical kindergarteners confirmed these findings (Apfelstadt, 1984; Welch et al., 1997). Results from a more recent study corroborated the finding that each of these task types were good discriminators of singing accuracy when all participants were asked to sing the same song (Nichols, 2013), and future studies should incorporate as many task types as is feasible and include enough items within each task type to assure internal reliability.

Presentation Variables

Contextual Presentation

Song singing can be described as contextualized because song presentation includes additional information that is absent in single-pitch-matching exercises. For example, songs provide cues like key, meter, and contour, which may support the developing student in the act of singing. This difference in context is important because it makes the two task types of pitch matching and song singing to be theoretically different and thus difficult to compare. No significant effect on the presence of accompaniment has been reported; thus, the role of accompaniment as a contextual cue is not considered here (Atterbury & Silcox, 1993; Guilbault, 2004).

Some researchers have designed their studies using contextually presented pitch-matching stimuli, as opposed to pitches presented in absence of a tonal center.

Downloaded from upd.sagepub.com by guest on April 17, 2015

Geringer (1983) established context for test items by asking children to sing back the final pitch in a simple three-measure song. Demorest and Clements (2007) adopted a similar procedure, and for comparison, they also included a single-pitch-matching condition that replicated the typical classroom exercise in which the teacher sings a single note and the student attempts to sing it back. Singers classified as certain and uncertain singers differed from one another, but did not differ significantly in their responses to the two conditions. Those singers classified as inconsistent (neither certain or uncertain), however, did differ. These inconsistent singers were between the certain and uncertain groups in terms of pitch-matching ability. The authors reported that if only the contextual condition were presented, the singers in this group would have been classified as good singers.

Demorest (2001) found that unlike inconsistent singers, certain and uncertain singers performed differently on a measure of pitch discrimination, which could suggest that pitch-discrimination ability plays a role in accurate pitch production. Geringer (1983) found no significant differences in pitch discrimination based on children's pitch-matching ability, though the tasks were relatively simple. Demorest and Clements (2007) reported that the results in these studies may vary due to the relative difficulty of items used, but that most studies have not directly compared tasks. They recommended further examination of task-based variability in singing accuracy (see Figure 1, Singing Response Modes).

Model Characteristics

Yarbrough et al. (1991) found that children responded more accurately to a female model than a male model. Children may also respond better to another child's voice, rather than an adult female's voice or an adult male's voice (Green, 1990). Overall, children sing more accurately when less vibrato is presented in the vocal model (Yarbrough et al., 1992). For classroom applications rather than research use, the range of potential vocal models varies, as do the teachers and instruments. For example, male timbre, falsetto, and sine wave models have been compared to demonstrate that male stimuli may be more effective than sine waves for both boys and girls (Price et al., 1994). Overall, girls responded better to higher stimuli, whereas boys responded better to lower stimuli. Additionally, the octave of the stimuli affected the octave of the response. Boys may need help in understanding whether they are expected to sing at, or below, a high pitch given from a female model (i.e., octave displacement).

Male teachers must carefully choose in what octave they sing to best help their students (Price et al., 1994; Yarbrough, Morrison, Karrick, & Dunn, 1995). Otherwise, teachers have few choices: They can use their own voice, use their students as models for one another, or use a common instrument like the piano, knowing that these choices may affect singing measures. Generally, children respond best when the stimulus is presented in their register (Kramer, 1986; Sims, Moore, & Kuhn, 1982). For singing assessment in classrooms, all of these options exist as possibilities for test stimuli.

Range

Children tend to select lower, rather than higher keys when singing, but when given a higher register, children can evidence wide ranges (Moore, 1991). Still, children often modulate keys to use pitches that may be easy or comfortable for them, and children may show greater ranges when echoing pitch-matching tasks than when singing songs from memory (Flowers & Dunne-Sousa, 1990). Rutkowski (1990) described range ability in terms of singing voice development, from a student with no use of the singing voice to a student with full use of an expanded range: (a) does not sing but chants text; (b) sustains tones with sensitivity to pitch, usually A2 to C3; (c) wavers between speaking/singing, usually up to F3; (d) has initial singing range, D3 to A3; and (e) has full singing range, beyond register lift near B-flat-3 and beyond. Rutkowski (1996) then added levels between these five intervals to describe the developing singing voice more specifically. Wolf (2005) and Nichols (2013) reported that low-range test items are easier than high ones, which corroborates Rutkowski's suggestion that range develops upward.

Text

Song lyrics could be theorized to add complexity to the task demands in singing and may take priority over accurate singing. For example, the first level of the Singing Voice Development Measure (Rutkowski, 1996) states that a student " . . . does not sing but chants song text." The beginning level of the scale developed by Welch et al. (1995) for assessment of singing accuracy also focuses on text rather than melody. Goetze (1985) found the neutral syllable loo to be easier to sing than text, and more so for the kindergarteners and first graders than the second and third graders in her sample. Smale (1987) also compared the use of text with the syllable loo but found no difference in performance using a sample of preschoolers. Gault (2002) reinforced the idea that lyrics must be considered in the context of the specific song and the varying levels of student ability and aptitude. In sum, the findings related to text versus neutral syllable must be considered as one of many assessment variables.

Downloaded from upd.sagepub.com by guest on April 17, 2015

Conclusions

Teachers are urged to consider using classroom assessments to evaluate students' singing voice development, range, ability to maintain a tonal center in increasingly complex passages, and matching intervals and patterns in increasing levels of difficulty. A student who does not have access to his head voice will demonstrate an inability to match pitches in the upper range. Thus, teachers should start the singing assessment trajectory by evaluating singing voice development, which is similar for developing singers at any age level. Since Rutkowski's (1996) scale has been deemed valid and reliable for certain grades, teachers already have access to a scale designed for this purpose.

Next, students can be asked to sing back phrases and short songs within their range while maintaining the key center. Teachers should exercise caution in advancing students too quickly through song repertoire. Instead, students can be asked to sing simple songs with varying tempi and in varying keys before proceeding to songs of greater length with more difficult intervals. Large interval leaps provide an opportunity for students to modulate to lower, more comfortable key centers, and teachers are encouraged to sequence song material to suit the individual singer's needs.

Last, students can be asked to match intervals and patterns that increase in difficulty by using the pitch sequences mentioned earlier by Moore et al. (1995), Jones (1971), Sinor (1984), Wolf (2005), and Young (1971). Teachers can use songs from textbooks or Kodaly and Orff Schulwerk methods to evaluate singers' abilities with specific pitch sequences. Singing research has employed pitch-matching tasks such as single pitches, intervals, patterns, phrase singing, and song-singing tasks. Roberts and Davies (1975) and Nichols (2013) established that at least four of these—the single pitch, interval, pattern, and song-singing tasks—are discriminators of accurate singing. Therefore, it is clear that these four tasks can be used in informal or formal singing assessments.

The studies reviewed here can be used to suggest several important considerations in designing an assessment of singing accuracy for classroom use. First, the task should be chosen based on the purpose for assessment. When a global picture of singing accuracy is desired, more than one task type should be used, especially as all task types have been shown to be useful discriminators of singing accuracy. Second, the construct of singing voice development must be taken into consideration since tasks requiring greater voice development can obscure the results on singing accuracy. Third, the variables of contextual presentation, model characteristics, range, and text must all be considered as critical aspects of test construction.

For formal assessment, teachers must carefully construct a singing test so the results represent their students in a valid and reliable way. As teacher and student accountability is increasingly expected in local schools, teachers must approach testing in music performance as carefully as their colleagues in other subject areas. A sample singing assessment cycle is available online with task and rubric examples (Supplemental material available at http://upd.sagepub.com/content/by/supplemental-data). Teachers can begin by assessing students informally using the progression of intervals suggested by researchers in the previous sections. Summative assessments should include pitch patterns and song-singing tasks. For researchers, the literature indicates a need for test construction that allows for comparison with other studies. Specifically, researchers might consider using measures that incorporate multiple task types, keeping in mind that performance on one song is difficult to compare with performance on a different song. More research in singing is needed on the task and presentation variables reviewed in this article to support the teaching and learning of this important skill.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Apfelstadt, H. (1984). Effects of melodic perception instruction on pitch discrimination and vocal accuracy of kindergarten children. *Journal of Research in Music Education, 32*, 15–24. doi:10.2307/3345277

Atterbury, B. W., & Silcox, L. (1993). The effect of piano accompaniment of kindergartners' developmental singing ability. *Journal of Research in Music Education, 41*, 40–47. doi:10.2307.3345478

Berkowska, M., & Dalla Bella, S. (2009). Acquired and congenital disorders of sung performance: A review. *Advances in Cognitive Psychology, 5*, 69–83. doi:10.2478/v10053-008-0068-2

Boardman, E. L. (1964). *An investigation of the effect of pre-school training on the development of vocal accuracy in young children* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6408354)

Brophy, T. S. (1997). Authentic assessment of vocal pitch accuracy in first through third grade children. *Contributions to Music Education, 24*(1), 57–70.

Cooper, N. A. (1995). Children's singing accuracy as a function of grade level, gender, and individual versus unison singing. *Journal of Research in Music Education, 43*, 222–231. doi:10.2307/3345637

Downloaded from upd.sagepub.com by guest on April 17, 2015

Dalla Bella, S., & Berkowska, M. (2009). Singing proficiency in the majority. *Annals of the New York Academy of Sciences, 1169*, 99–107. doi:10.1111/j.1749-6632.2009.04558.x

Dalla Bella, S., Giguère, J. F., & Peretz, I. (2007). Singing proficiency in the general population. *Journal of the Acoustical Society of America, 121*, 1182–1189. doi:10.1121/1.2427111

Demorest, S. M. (2001). Pitch matching performance of junior high boys: A comparison of perception and production. *Bulletin of the Council for Research in Music Education, 151*, 63–70.

Demorest, S. M., & Clements, A. (2007). Factors influencing the pitch matching of junior high boys. *Journal of Research in Music Education, 55*, 190–203. doi:10.1177/002242940705500302

Flowers, P. J., & Dunne-Sousa, D. (1990). Pitch-pattern accuracy, tonality, and vocal range in preschool children's singing. *Journal of Research in Music Education, 38*, 102–114. doi:10/2307/3344930

Gault, B. (2002). Effects of pedagogical approach, presence/absence of text, and developmental music aptitude on the song performance accuracy of kindergarten and first-grade students. *Bulletin of the Council for Research in Music Education, 152*, 54–63.

Geringer, J. (1983). The relationship of pitch matching and pitch-discrimination abilities of preschool and fourth-grade students. *Journal of Research in Music Education, 31*, 93–99. doi:10.2307/3345213

Goetze, M. (1985). *Factors affecting accuracy in children's singing* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8528488)

Goetze, M., & Cooper, N., & Brown, C. (1990). Recent research on singing in the general music classroom. *Bulletin of the Council for Research in Music Education, 104*, 16–37.

Green, G. A. (1990). The effect of vocal modeling on pitch matching accuracy of elementary schoolchildren. *Journal of Research in Music Education, 38*, 225–231. doi:10.2307/3345186

Green, G. A. (1994). Unison versus individual singing and elementary students' vocal pitch accuracy. *Journal of Research in Music Education, 42*, 105–114. doi:10.2307/3345186

Guerrini, S. C. (2006). The developing singer: Comparing the singing accuracy of elementary students on three selected vocal tasks. *Bulletin of the Council for Research in Music Education, 167*, 21–31.

Guilbault, D. M. (2004). The effect of harmonic accompaniment on the tonal achievement and tonal improvisations of children in kindergarten and first grade. *Journal of Research in Music Education, 52*, 64–76. doi:10.2307/3345525

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. doi:10.1177/8755123312438516

Jones, M. (1971). A pilot study in the use of a vertically-arranged keyboard instrument with the uncertain singer. *Journal of Research in Music Education, 19*, 183–194. doi:10.2307/3343822

Joyner, D. R. (1969). The monotone problem. *Journal of Research in Music Education, 17*, 115–124. doi:10.2307/3344198

Klemish, J. (1974). Treating the uncertain singer through the use of the tape recorder. *Bulletin of the Council for Research in Music Education, 37*, 36–45.

Kramer, S. J. (1986). *The effects of two different music programs on third and fourth grade children's ability to match pitches vocally* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8524224)

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*, 35–47. doi:10.2307/3345758

McCoy, C. (1997). Factors relating to pitch matching skills of elementary education majors. *Journal of Research in Music Education, 45*, 356–366.

Moore, R., Fyk, J., Frega, A., & Brotons, M. (1995). Influences of culture, age, gender and two-tone melodies on interval matching skills of children from Argentina, Poland, Spain and the USA. *Bulletin of the Council for Research in Music Education, 127*, 127–135.

Moore, R. S. (1991). Comparison of children's and adults' vocal ranges and preferred tessituras in singing familiar songs. *Bulletin of the Council for Research in Music Education, 107*, 13–22.

Murry, T., & Zwirner, P. (1991). Pitch matching ability of experienced and inexperienced singers. *Journal of Voice, 5*, 197–202. doi:10.1016/S0892-1997(05)80187-0

Muse, M. B. (1993). *A comparison of two methods of teaching singing to primary children: An attempt to determine which of two approaches to teaching singing is more effective* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 1358052)

Nichols, B. E. (2013). *Task-based variability in children's singing accuracy* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3588972)

Petzold, R. G. (1963). The development of auditory perception of musical sounds by children in the first six grades. *Journal of Research in Music Education, 11*, 21–43. doi:10.2307/3344529

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

Pfordresher, P. Q., Brown, S., Meier, K. M., Belyk, M., & Liotti, M. (2010). Imprecise singing is widespread. *Journal of the Acoustical Society of America, 128*, 2182–2190. doi:10.1121/1.3478782

Phillips, K., & Aitchison, R. (1997). Effects of psychomotor instruction on elementary general music students' singing performance. *Journal of Research in Music Education, 45*, 185–196.

Phillips, K., & Aitchison, R. (1999). Second-year results of a longitudinal study of the relationship of singing instruction, pitch accuracy, and gender to aural acuity, vocal achievement, musical knowledge, and attitude towards singing among general music students. *Contributions to Music Education, 26*(1), 67–85.

Phillips, K., Aitchison, R., & Nompula, Y. (2002). The relationship of music aptitude to singing achievement among fifth grade students. *Contributions to Music Education, 29*(2), 47–58.

Phillips, K., & Doneski, S. (2011). Research on elementary and secondary school singing. In R. Colwell & P. Webster

Downloaded from upd.sagepub.com by guest on April 17, 2015

(Eds.), *MENC handbook of research on music learning, Volume 2: Applications* (pp. 176–232). New York, NY: Oxford University Press.

Phillips, K., & Vispoel, W. (1990). The effects of class voice and breath-management instruction of vocal knowledge, attitudes, and vocal performance among elementary education majors. *Quarterly Journal of Music Teaching and Learning, 1*, 96–105.

Porter, S. Y. (1977). The effect of multiple discrimination training on pitch matching behaviors of uncertain singers. *Journal of Research in Music Education, 25*, 68–82. doi:10.2307/3344846

Price, H. E., Yarbrough, C., Jones, M., & Moore, R. S. (1994). Effects of male timbre, falsetto, and sine-wave models on interval matching by inaccurate singers. *Journal of Research in Music Education, 42*, 269–284. doi:10.2307/3345736

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

Rutkowski, J. (1990). The measurement and evaluation of children's singing voice development. *Quarterly Journal of Music Teaching and Learning, 1*(1–2), 81–95.

Rutkowski, J. (1996). The effectiveness of individual/small-group singing activities on kindergartners' use of singing voice and developmental music aptitude. *Journal of Research in Music Education, 44*, 353–368. doi:10.2307/3345447

Rutkowski, J., & Miller, M. S. (2003). The effect of teacher feedback and modeling on first graders' use of singing voice and developmental music aptitude. *Bulletin of the Council for Research in Music Education, 156*, 1–10.

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. doi:10.1177/8755123310378454

Sims, W. L., Moore, R. S., & Kuhn, T. L. (1982). Effects of female and male vocal stimuli, tonal pattern length and age on vocal pitch-matching abilities of young children from England and the United States. *Psychology of Music* (Special issue), 104–108.

Sinor, E. (1984). *The singing of selected tonal patterns by preschool children* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8501456)

Smale, M. J. (1987). *An investigation of pitch accuracy of four- and five-year-old singers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8723851)

Smith, R. (1973). *Factors related to children's in-tune singing abilities* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7411404)

Van Zee, N. (1984). An investigation study of young children's vocal problems and remedial needs. *Missouri Journal of Research in Music Education, 5*, 55–71.

Watts, C., Moore, R., & McCaghren, K. (2005). The relationship between vocal pitch matching skills and pitch discrimination skills in untrained accurate and inaccurate singers. *Journal of Voice, 19*, 534–543. doi:10.1016/j.jvoice.2004.09.001

Welch, G., Sergeant, D., & White, P. J. (1995). The singing competencies of five-year-old developing singers. *Bulletin of the Council for Research in Music Education, 127*, 155–162.

Welch, G., Sergeant, D. C., & White, P. J. (1997). Age, sex, and vocal task as factors in singing "in tune" during the first years of schooling. *Bulletin of the Council for Research in Music Education, 133*, 153–160.

Western, B. (2002). *Fundamental frequency and pitch matching accuracy characteristics of first grade general music students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3073408)

Wise, K. J., & Sloboda, J. A. (2008). Establishing an empirical profile of self-defined "tone deafness": Perception, singing performance and self-assessment. *Musicae Scientiae, 12*(1), 3–26. doi:10.1177/102986490801200102

Wolf, D. (2005). A hierarchy of tonal performance patterns for children ages five to eight years in kindergarten and primary grades. *Bulletin of the Council for Research in Music Education, 163*, 61–68.

Wurgler, P. S. (1990). *A perceptual study of vocal registers in the singing voices of children* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8501456)

Yarbrough, C., Bowers, J., & Benson, W. (1992). The effect of vibrato on the pitch matching accuracy of certain and uncertain singers. *Journal of Research in Music Education, 40*, 30–38. doi:10.2307/3345772

Yarbrough, C., Green, G., Benson, W., & Bowers, J. (1991). Inaccurate singers: An exploratory study of variables affecting pitch matching. *Bulletin of the Council for Research in Music Education, 107*, 23–34.

Yarbrough, C., Morrison, S. J., Karrick, B., & Dunn, D. E. (1995). The effect of male falsetto on the pitch-matching accuracy of uncertain boy singers, grades K-8. *Update: Applications of Research in Music Education, 14*(1), 4–10.

Young, W. (1971). *An investigation of the singing abilities of kindergarten and first grade children in east Texas* (ERIC EDO 69431). Retrieved from http://files.eric.ed.gov/fulltext/ED069431.pdf

Downloaded from upd.sagepub.com by guest on April 17, 2015