

# Effect of vocal versus piano doubling on children's singing accuracy

Psychology of Music

1–9

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0305735620936757

journals.sagepub.com/home/pom

Bryan E Nichols 

## Abstract

Children's singing development is an important part of the music classroom, where instruction is often assisted by the teacher's voice or the piano. However, it is unknown whether children sing more accurately when doubled by another voice or instruments. The purpose of this study was to investigate the effect of doubling timbre on children's singing accuracy. Third- and fourth-grade children ( $N = 61$ ) performed pitch matching and song singing tasks doubled by pre-recorded vocal and piano stimuli, counterbalanced to control for order effects. Performance when doubled by voice and doubled by piano was significantly and strongly correlated,  $r(59) = .81$ ,  $p < .001$ . Children performed more accurately on the vocal doubling condition than the piano ( $p = .002$ ) on pitch matching tasks, but there was no significant difference on the song singing task.

## Keywords

*assessment, child development, education, pedagogy, pitch, singing, timbre*

Music educators work to develop children's voices throughout the elementary school years by providing opportunities to sing many songs, to develop good singing tone, and to exercise their upper singing range. The last of these—expanding the range—is referred to as singing voice development (Rutkowski, 1990) and is an important component of accurate singing: children must be able to access all parts of the vocal range to sing well. If children can access a certain part of the range, they are more likely to sing successfully in tune. Effective instruction can lead children to sing in tune, a construct commonly referred to as vocal pitch accuracy or singing accuracy (Demorest, Nichols, & Pfordresher, 2018). Assisting children's singing skills is an important part of classroom teaching, and instruction is often assisted by the teacher's voice, other students' voices, or commonly a piano. When the external stimulus plays the singers' part simultaneously, this is sometimes referred to as *doubling*, a term that will be used here to describe an external stimulus playing the singer's part exactly.

School of Music, Penn State University, University Park, PA, USA

## Corresponding author:

Bryan E Nichols, School of Music, Penn State University, Music Building, University Park, PA 16802, USA.

Email: bnichols@psu.edu

Singing instruction has been shown to improve singing accuracy (Demorest et al., 2018), which is important because singing accuracy seems to decline after developing in the elementary school years (Demorest & Pfordresher, 2015). Recent developments in computerized testing (Cohen, 2015; Demorest et al., 2015) and scoring procedures (Larrouy-Maestri, Leveque, Schon, Giovanni, & Morsomme, 2013; Pfordresher & Larrouy-Maestri, 2015) have offered teachers the promise of reliable and valid assessments so progress in singing accuracy can be more reliably tracked. Broad adoption of simpler assessments may help to improve instruction. Furthermore, teacher accountability in schools requires such assessments (Salvador, 2010), and work in this area has encouraged instruction, assessment, and retention in music ensembles.

Previous studies of college students have demonstrated superior performance in a doubling condition, leading to a belief that singers and instrumentalists are supported by an external stimulus (Geringer, 1978; Vorce, 1964). However, the effectiveness of pedagogical techniques may differ when applied to young singers rather than adult musicians, and, in fact, anecdotal evidence supports the notion that some children may perform more accurately singing in a solo singing condition (Goetze & Horii, 1989). In one study of a teacher-identified “monotone” sample of children, researchers found a subset of the individuals to sing more in-tune when isolated from the others (Roberts & Davies, 1975).

Previous studies specifically comparing solo to doubled singing conditions in children indicate mixed results for singing accuracy. Some authors report superior singing in the solo singing condition (Goetze & Horii, 1989; Smale, 1987), while others report superior doubled singing (Green, 1994; Nichols, 2016b) or no significant difference (Cooper, 1995; Smith, 1973). These previous attempts at clarifying the relationship between the two singing response modes used prevailing methodologies such as the use of age-appropriate songs or pitch matching tasks in conjunction with rating scales or acoustic scoring in cent deviation (Nichols, 2016a).

Adolescents vary in response to a female vocal model, male vocal model, or synthesized accompaniment depending on their voice change status (Oberfield, 2005), but it is unknown how children respond to an instrument such as piano compared to a voice for modeling. Women aged 20 to 30 matched their own voice better than a female vocal model or a synthesized tone (Moore, Estis, Gordon-Hickey, & Watts, 2008). For solo singing, children sing more accurately when presented with stimuli in their register (Kramer, 1986; Sims, Moore, & Kuhn, 1982), if a female vocal model is used (Yarbrough, Green, Benson, & Bowers, 1991), when a child’s voice is used (Green, 1990), lesser vibrato is used (Yarbrough, Bowers, & Benson, 1992), and if presented a male falsetto (rather than a chest voice; for example, Price, Yarbrough, Jones, & Moore, 1994; Yarbrough, Morrison, Karrick, & Dunn, 1995). Children grades K-8 respond more accurately to vocal models than to a sine-wave model, though sine-wave tones are not representative of piano usage in classroom settings (Price et al., 1994). Whether piano accompaniment is used for children’s singing instruction did not influence singing ability in Kindergartners assigned to accompaniment and no accompaniment conditions during a year of instruction (Atterbury & Silcox, 1993). Guilbault (2004) confirmed harmonic accompaniment did not significantly affect children’s song performance, but the effect of piano timbre for doubling the voice part is unknown.

A recent analysis of the previous literature on doubling by other voices presented evidence to explain the differences in previous research (Nichols & Lorah, 2019). The individual performance among study participants determines the degree of difference between overall solo performance and overall doubled performance, possibly accounting for the mixed findings in previous studies as to the singing accuracy in solo or doubled conditions. Some task types or singing conditions may elicit favorable doubled singing among certain participants. Goetze and

Horii (1989) suggested weak singers perform more accurately alone than they do singing along with their peers. Furthermore, the difficulty of specific singing test items may influence the outcomes in previous studies, showing some singer types to elicit superior solo performance on easier tasks or when stimuli are presented using certain methods. No systematic explanation for the contrasting previous results has been established.

Beyond these explanations, external, ecological factors such as teacher training, curriculum differences, or regional differences may affect the proportion of students who sing better in the solo condition versus those who sing better in the doubled condition. Thus, further work exploring the discrepancy in these previous studies is warranted. The purpose of this study is to further explore doubling effects in singing accuracy, specifically, to explore the effect of doubling timbre on singing accuracy in a commonly used test construction. The research question was as follows:

**Research Question 1:** Does doubling timbre affect singing accuracy?

## Method

### *Participants*

Third- and fourth-grade children ( $M_{\text{age}} = 9$  years, 8 months) were recruited during music class at a local suburban public elementary school in the Midwest United States using approved Institutional Review Board (IRB) procedures. Child assent and parental consent protocols were followed, and participants were tested one-on-one by a research assistant during their normal weekly music period in a room near the music classroom. Participants were not prescreened for ability or on any other factors. The school was a suburban public elementary school served by one music teacher who met intact K–fourth-grade classes twice weekly.

### *Test design*

The test compared singing accuracy in two conditions: doubled by a vocal model or by a piano timbre. In a previous study, evidence suggested children can be reliably assessed on pitch matching tasks using a minimum of three of the five items presented per pitch matching task (Nichols, 2016b). Thus, we revised the previous test, shortening it to three single pitch items, three intervals, and three four-note patterns, and identical piano stimuli were created using notation software audio export (Finale 25) to match the previously recorded vocal stimuli. The same pitch sequences were used for the vocal and piano doubling conditions. The song singing task of *Jingle Bells* was used after the pitch matching items in each condition. One presentation of the single pitch was used (e.g., D) and one presentation of each pitch in the interval was used (e.g., A–G).

### *Procedure*

Stimuli were played for participants using stereo headphones, monitored by a research assistant also using headphones, for a total test session lasting less than 10 min for each child. The test included pitch presentations at 60 bpm on the neutral syllable /du/ based on previous research (Sinor, 1984; Wolf, 2005) in the range of D4 to A4 (per Wolf, 2005) but also to avoid a confound with singing development in the range above the B-flat register transition (Rutkowski, 1990). Children may be reliably assessed on the first attempt, so only one attempt

at each item was provided (Nichols & Wang, 2016). Pitches were presented from a laptop computer, then echoed by the participant as the stimuli played a second time for the doubling stimulus. Thus, the doubling stimulus was identical to the modeling stimulus. Finally, the test concluded with a song singing task using *Jingle Bells* presented in the same key center (*D*) as the pitch matching tasks. For this task, the children were given the starting pitch on the third scale degree (*F-sharp*) and heard a pre-recorded voice, "Ready, set, sing." The participant did not hear a modeling of the song; thus, the pre-recorded stimulus for the song singing task was only presented during performance. Approximately half the participants begin in the vocal doubling condition and half began in the piano doubling condition.

### Stimuli

Previously used vocal stimuli were used for this study. The vocal stimuli were recorded by a female model tuning to A (440Hz) instructed to employ minimal vibrato (Yarbrough et al., 1991). The piano stimuli were notated in software and exported using a "Grand Piano" timbre set for playback at 60 bpm, then exported as mp3 files for use as stimuli.

### Scoring

Pitch matching items were analyzed in cent deviation using the middle, stable portion of the participant's recorded samples from the test. Deviation scores were calculated in Hertz from the actual pitch stimuli the participants heard and then transformed to cents for analysis. Using Praat software, the middle stable portion of each pitch was selected and analyzed by a research assistant, and 10% of samples were also analyzed by the researcher for confirmation. Deviation scores for individual pitches are negative if the participant sang below the stimulus pitch (flat), and deviation scores are positive if the participant sang above the stimulus pitch (sharp). Scores were transformed to absolute value deviation scores (unsigned deviation scores) and then transformed to unsigned deviation scores in cents for the statistical comparisons.

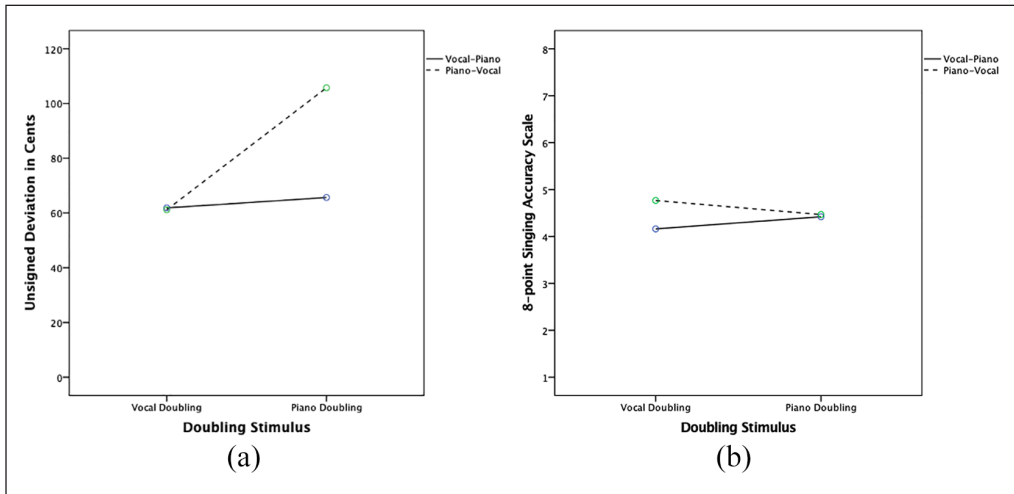
For song singing, two raters with expertise in singing instruction scored a randomly selected 10% of the participants with acceptable inter-rater reliability ( $r = .87$ ) using a common song singing scale shown in Figure 1 (Wise & Sloboda, 2008). One rater scored the remaining participants. The raters were graduate students with expertise in teaching vocal music to children; they were blind to the child's gender, home classroom, and academic status in school. One participant indicated a history of hearing impairment; this participant performed within 1 *SD* of the mean on all tasks except one, and we chose to keep the participant in the sample for analysis (see Figure 2).

### Results

Third- ( $n = 33$ ) and fourth-grade ( $n = 28$ ) children ( $N = 61$ ) completed singing test in two forms counterbalanced for order effect; Form A: vocal doubling condition followed by piano doubling condition ( $n = 31$ ), and Form B: piano doubling condition followed by vocal doubling condition ( $n = 30$ ). Performance in the two timbre conditions for pitch matching (voice and piano) was significantly and strongly correlated,  $r(59) = .89$ ,  $p < .001$ . To analyze pitch matching data, a repeated-measures analysis of variance (ANOVA) with Form (whether the participant began the test with vocal doubling or piano doubling first) as a between-subjects

8 – All melody is accurate and in tune, and key is maintained throughout.
7 – Key is maintained throughout, and melody accurately represented, but some mistunings (though not enough to alter the pitch-class of the note).
6 – Key is maintained throughout and melody mostly accurately represented, but some errors (notes mistuned sufficiently to be ‘wrong’).
5 – Melody largely accurate, but singer’s key drifts or wanders. This may be the result of a mistuned interval, from which the singer then continues with more accurate intervals but without returning to the original pitch.
4 – Melody fairly accurate, or mostly accurate within individual phrases, but singer changes key abruptly, especially between phrases (e.g. adjusting higher-lying phrases down).
3 – Singer accurately represents the contour of the melody but without consistent pitch accuracy or key stability.
2 – Words are correct but there are contour errors. Pitches may sound almost random.
1 – Singer sings with little variation in pitch, and may chant in speaking voice rather than singing.

**Figure 1.** Singing Accuracy Scale from Wise and Sloboda (2008).



**Figure 2.** (a) Pitch Matching and Song Singing Mean by Doubling Stimulus (Lower Deviation Score Indicates Higher Accuracy). (b) Song Singing Mean by Doubling Stimulus (Higher Score Indicates Higher Accuracy).

factor indicated a significant difference between singing accuracy when doubled by a vocal or a piano model,  $F(1, 118) = 10.07, p = .002, \eta_p^2 = .146$  (see Table 1). Participants sang pitch matching tasks more accurately in the vocal doubling condition than the piano doubling condition. There was a significant Doubling Timbre by Form interaction,  $F(1, 18) = 7.17, p = .01, \eta_p^2 = .108$  (see Figure 1), indicating those who sang the vocal doubling condition first performed the piano condition more accurately than those who began with piano doubling. For the song singing task on *Jingle Bells*, there was no significant difference between the two doubling conditions,  $p > .05$ .

**Table 1.** Descriptive statistics for pitch matching deviation scores ( $N=61$ ).

	Doubling stimulus	Mean deviation <sup>a</sup>	SD
Single pitch	Vocal	52.7	72.5
Interval	Vocal	60.5	54.9
Pattern	Vocal	71.4	56.3
Single pitch	Piano	80.0	118.2
Interval	Piano	79.9	107.9
Pattern	Piano	96.3	107.6

<sup>a</sup>Unsigned deviation scores in cents (lower is more accurate).

## Discussion

The findings from this study suggest that doubling timbre affects children's singing accuracy in a comparison of vocal to piano doubling conditions. Children performed pitch matching items more accurately in the vocal doubling condition, which presents ecological validity challenges to future testing for vocal versus instrumental doubling: children may have found it more natural to sing along to a pre-recorded voice than the piano timbre and thus were more accurate. For participants assigned to perform the vocal doubling condition first, vocal doubling may have increased performance in the piano condition which followed. Most importantly, this interaction where participants who performed with vocal doubling first indicated more accurate performance with piano doubling has implications for testing in singing research and perhaps also for music classrooms.

Previous research suggests poor singers may be overwhelmed by the presence of external stimuli such as the doubling condition used in this study (Cooper, 1995; Smith, 1973); those singers may perform more accurately alone, and this "type" of singer may need to be categorized formally as different from other singer types. Some singers performed accurately in one condition and similarly in the other, while others performed accurately in one condition but with low accuracy in another condition. Perhaps this presents evidence that the proportion of singer types (those who sing accurately in certain conditions or on certain singing tasks) varies in student populations and moderates the effect of doubling timbre. There was no significant difference in song singing performance (while there was in pitch matching performance) between the two doubling conditions, but performance may have varied if a different song task were used such as an easier or more difficult song, or a newly taught song instead of a familiar song. Children's performance on newly taught song material in various doubling conditions is still unknown, and this presents an opportunity for future research.

Due to the range of possibilities in teacher and student characteristics, and the unpredictable difficulty of pitch sequences and song tasks for any group of students, a discussion of limitations is presented in the following paragraphs. First, it must be noted that some elementary music teachers may do mostly a cappella singing instruction, while others may use keyboard or other accompaniment such as guitar. The results of doubling studies must certainly vary between participants from different types of these instructional conditions. The music classroom from which participants were sampled in this study did not have a focus on piano-assisted instruction (teacher interview).

Second, some natural cues may come from a vocal stimulus—such as audible breath inhalation or consonant formations—that may assist some or all singers in doubling that do not exist for instrumental doubling. Thus, smaller differences may exist between the doubling effect of

different instrument timbres than may exist than between vocal and instrumental conditions, a proposition not answered with the current data comparing vocal to piano doubling. In addition, visual cues such as a head nod or conducting gesture may help students in the classroom setting respond to instrumental doubling, whereas doubling in this study was provided by a pre-recorded audio track with no visual stimuli. Future studies could replicate the visual stimuli to approximate the ecological validity of the teacher presence in classroom instrumental doubling.

Next, previous evidence in favor of task-based variability may indicate that certain pitch matching task types are easier or harder than others, or even that individual items within task types were of greater or lesser difficulty. Or, the relationship to those task types and specific items may vary in similarity to the tasks and specific pitch sequences students practice in the classroom. The difficulty indices of specific intervals have been studied, but there is less data on the difficulty and discrimination indices for the multitude of pitch combinations and pattern lengths in longer pitch pattern items such as the four-note patterns used here.

Finally, Goetze and Horii (1989) suggested poor singers perform more accurately alone than they do singing along with their peers. The previous mixed findings as to the superior performance in solo or doubled singing, and the present findings of no difference between the two conditions may be due in part to treating overall accurate and overall inaccurate singer types as one group when the opposing performance of the two singer types may negate the actual singing effects based on factors like doubling. Future research might include comparisons to solo singing, accounting for test length for young singers which limited the test construction for the present study. However, Pfordresher and Larrouy-Maestri (2015) warned against dichotomizing singer types for statistical comparison, though they suggest it depends on the specific research question. For investigations of children with developing voices and musical abilities—and in education settings generally—it may be useful to form remedial versus normative groupings (e.g., the common red bird vs. blue jay reader groups within a class). Furthermore, it may be useful for teaching or for research purposes to define categories of singer types—if they exist—of children who sing more accurately alone (on certain tasks) versus children who sing more accurately doubled under certain conditions.

Performance on pitch matching versus song singing tasks was highly correlated, suggesting some future potential predictive validity to the use of pitch matching tasks to evaluate developing singers' song singing skill. This relationship between tasks has been previously reported (Demorest et al., 2018) and can be explored in future research. The nature of call and response tasks such as those used in the pitch matching test versus recall tasks such as the song task used may moderate the effect of the difficulty levels presented by specific pitch matching and song singing items (see Nichols, 2016a, for an outline of task types). How would performance have differed if an unfamiliar song were presented for evaluation in conditions, including (1) solo singing, (2) vocal doubling, and (3) piano doubling?

Finally, it is not surprising that participants performed more accurately in the vocal doubling versus piano doubling condition: singers are shown to perform more accurately in response to vocal model stimuli and are here shown to respond more accurately to vocal doubling. A previous comparison of subjective versus objective (acoustic) scoring methods yields highly related results (Larrouy-Maestri et al., 2013) supporting the conclusions we present using acoustic measures for pitch matching and expert raters for song singing. However, additional research should be undertaken to confirm these results in children and to support the standardized testing protocols employed in the present study such as the Seattle Singing Accuracy Protocol (SSAP, Demorest et al., 2015) and Advanced Interdisciplinary Research in Singing (AIRS) Test Battery (Cohen, 2015).

## Funding

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

## ORCID iD

Bryan E Nichols  <https://orcid.org/0000-0002-8970-3284>

## References

- Atterbury, B. W., & Silcox, L. (1993). The effect of piano accompaniment on Kindergarteners' developmental singing ability. *Journal of Research in Music Education*, 41, 40–47. doi:10.2307/3345478
- Cohen, A. (2015). The AIRS test battery of singing skills: Rationale, item types, and lifetime scope. *Musicae Scientiae*, 19, 238–264. doi:10.1177/1029864915599599
- 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
- Demorest, S. M., Nichols, B. E., & Pfordresher, P. (2018). The effect of focused instruction on young children's singing accuracy. *Psychology of Music*, 46, 488–499. doi:10.1177/0305735617713120
- Demorest, S. M., & Pfordresher, P. Q. (2015). Singing accuracy development from K-adult: A comparative study. *Music Perception*, 32, 292–302. doi:10.1525/MP.2015.32.3.293
- Demorest, S. M., Pfordresher, P. Q., Bella, D. S., Hutchins, S., Loui, P., Rutkowski, J., & Welch, G. (2015). Methodological perspectives on singing accuracy: An introduction to the special issue on singing accuracy (Part 2). *Music Perception*, 32, 266–271. doi:10.1525/mp.2015.32.3.266
- Geringer, J. M. (1978). Intonational performance and perception of ascending scales. *Journal of Research in Music Education*, 26, 32–40. doi:10.2307/3344787
- Goetze, M., & Horii, Y. (1989). A comparison of the pitch accuracy of group and individual singing in young children. *Bulletin of the Council for Research in Music Education*, 99, 57–73.
- Green, G. A. (1990). The effect of vocal modeling on pitch matching accuracy of elementary schoolchildren. *Journal of Research in Music Education*, 38(3), 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
- 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
- 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 Digital Dissertations. (AAT 8524224)
- Larrouy-Maestri, P., Leveque, Y., Schon, D., Giovanni, A., & Morsomme, D. (2013). The evaluation of singing voice accuracy: A comparison between subjective and objective methods. *Journal of Voice*, 27, 251–259. doi:10.1016/j.jvoice.2012.11.003
- Moore, R. E., Estis, J., Gordon-Hickey, S., & Watts, C. (2008). Pitch discrimination and pitch matching abilities with vocal and nonvocal stimuli. *Journal of Voice*, 22, 399–407. doi:10.1016/j.jvoice.2006.10.013
- Nichols, B. E. (2016a). Critical variables in singing accuracy test construction: A review of literature. *Update: Applications of Research in Music Education*, 35, 39–46. doi:10.1177/8755123315576764
- Nichols, B. E. (2016b). Task-based variability in children's singing accuracy. *Journal of Research in Music Education*, 64, 309–321. doi:10.1177/0022429416666054
- Nichols, B. E., & Lorah, J. (2019). Does doubled singing increase children's singing accuracy? A re-examination of previous findings. *Psychology of Music*, 48, 315–324. doi:10.1177/0305735618799171
- Nichols, B. E., & Wang, S. (2016). The effect of repeated attempts and test-retest reliability in children's singing accuracy. *Musicae Scientiae*, 20, 551–562. doi:10.1177/1029864916638914
- Oberfield, M. L. (2005). *Effect of homogeneous versus heterogeneous vocal stimulus on the simultaneous pitch-matching accuracy of middle school music students* (Master's thesis). Available from ProQuest Dissertations & Theses. (EP31793)



- Pfordresher, P. Q., & Larrouy-Maestri, P. (2015). On drawing a line through the spectrogram: How do we understand deficits of vocal pitch imitation? *Frontiers in Human Neuroscience*, *9*, Article 271. doi:10.3389/fnhum.2015.00271
- 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. *The Quarterly Journal of Music Teaching and Learning*, *1*, 81–95.
- 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*, 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, Spec Iss*, 104–108.
- Sinor, E. (1984). *The singing of selected tonal patterns by preschool children* (Doctoral dissertation). Available from ProQuest Digital Dissertations. (AAT 8501456)
- Smale, M. J. (1987). *An investigation of pitch accuracy of four- and five-year-old singers* (Digital dissertation). Available from ProQuest Digital Dissertations. (AAT 8723851)
- Smith, R. (1973). *Factors related to children's in-tune singing abilities* (Doctoral dissertation). Available from ProQuest Digital Dissertations. (AAT 7411404)
- Vorce, F. W. (1964). *The effect of simultaneous stimulus on vocal pitch accuracy* (Doctoral dissertation). Available from ProQuest Dissertations & Theses. (AAT 6500323)
- 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*, 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.
- 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*, 4–10.