Article

National Association for Music Education

Update
1–10

DOI: 10.1177/87551233211040726
update.sagepub.com
SAGE

Starting Pitch Selection Is Precise in Exploratory Study of Collegiate Nonmusic Majors

Bryan E. Nichols[1] and Annie Liu[1]

Implications for Music Teaching and Learning

- Incorporating knowledge of a song's range, especially in relation to a singer's own vocal range, could help them learn to select appropriate starting pitches or keys whenever singing without a reference pitch.
- For singers whose only singing experiences have been in ensembles like choir, heightening their self-awareness of their vocal range and key considerations could encourage them to sing more comfortably and confidently, perhaps for their lifetime.
- These nonmajor collegiate singers began songs on higher starting pitches than nonmajors in past studies, indicating a wider use of vocal ranges.
- These participants sang in tune, and singing in choir may help individuals be more accurate in their pitch and contribute to wider comfortable range use.
- Choral music curriculum at elementary, secondary, and postsecondary levels may be adjusted so that singers learn about their individual voices including identifying their voice type, preferred voice part, and choosing a key when singing familiar or unfamiliar songs.

Abstract

Starting pitch selection—whether intentional or habitual or otherwise— has been theorized to vary according to multiple variables. The purpose of this study was to explore starting pitch selection in undergraduate nonmusic major singers by using four well known songs and nonsinging tasks. Participants were actively singing in a college choir and majoring in fields outside of music. Performances indicated very small deviations in the median starting pitch for each criterion song (100–200 cents). Data indicated discrimination in starting pitch selection between songs. Active singers may place songs in the middle portion of the vocal range compared with the tendency toward lower range in nonsingers in previous research. A history of voice lessons was significantly associated with a positive response to a follow-up questionnaire item "I chose an intentional range for my voice."

Keywords

key selection, music memory, pitch selection, singing accuracy, starting pitch

Songs like "Happy Birthday" and other folk, holiday, and ceremonial music are frequently sung by individuals or groups, often spontaneously, or led by a music conductor. In informal settings, the music serves a purpose other than concert-making: to celebrate a function other than the music itself (e.g., alma maters and national anthems). Still, communal singing might be enjoyed by both singers and listeners when the music is produced in time (coordinated tempo) and in-tune (accurate intervals or performed in the same key), possibly leading to increased music participation (Demorest et al., 2017). The latter of these constructs, singing in-tune, is termed "singing accuracy" in the literature on the tunefulness in singers (e.g., Nichols, 2016). For singing accuracy in informal or ceremonial singing, inaccurate singing can arise due to people recognizing but not having learned a song, depending on definitions of learning and knowing one (Killian, 1996). Thus, it is possible not every individual will have practiced singing familiar songs, or they may not have heard accurate singing models. Furthermore, difficulty singing these may be due to starting to sing on a pitch ill-suited for the singing range; successful performance has been

[1]The Pennsylvania State University, University Park, USA

Corresponding Author:
Bryan E. Nichols, The Pennsylvania State University, 233 Music Building, University Park, PA 16802, USA.
Email: bnichols@psu.edu

indicated by singing songs with smaller ranges or through careful selection of starting pitch, as seen in musicians who possess the ability to select keys appropriate for the singer (Cevasco, 2008). One outcome of vocal music education could be the ability to independently sing the familiar songs from a culture in a range suitable to the singer's voice.

The range of a song and one's own vocal range offer implications for singing familiar songs, as do other specific features of the music. For example, pitch sequences may vary in terms of difficulty, and each melody represents different challenges (Nichols, 2016). In children, song features such as modality (major vs. minor), melodic contour, and harmonic function of a pattern were not shown to determine accuracy in performance (Wolf, 2005). Rather, difficulty was determined by intervallic relationships, length of pattern, and range. If modality, melodic contour, and harmonic function do not play a significant role in singing accuracy, then a high level of theory knowledge may not be required for singers as those aspects generally require more advanced music theory. Thus, theory preparation may not have a large effect on performance accuracy but may still influence the ability to choose keys and ranges for individual performance.

Children and adults vary in terms of individual vocal range. Undergraduate education majors in a music fundamentals class indicated wide ranges of about two octaves on average (Kuhn et al., 1979). The authors described participants as those singing at pitch (treble clef voices) or those singing an octave lower (bass clef voices), and they combined these participants into one group for analysis. Presumably, treble clef and bass clef singers may be assumed to have similar ranges, displaced lower by one octave for typical tenor or bass voice types. The results indicated an average vocal range of E3 to D5. The authors claimed a range of G-flat3 to C5 would include 75% of those students, which offers a suggestion for what ranges adults might use for singing typical, familiar songs in a key of their choosing. This range was large enough to accommodate common tunes like "Happy Birthday" and "The Star-Spangled Banner," and according to Kuhn et al. (1979), the range was suitable for singing the repertoire common at the time of the study found in music fundamental books and basal series texts.

Undergraduate nonmusic majors enrolled in elementary music classes have demonstrated the ability to reliably self-assess the vocal range, including measures such as the highest comfortable pitch, the lowest comfortable pitch, and the range of comfortable pitches (Geringer et al., 1980). Children demonstrated small ranges in comparison to the adults and less reliable self-assessment; the adults indicated similar ranges to Kuhn et al. (1979). Geringer et al. did not report significant differences between self-assessed range, teacher-assessed range, and importantly, teacher-reinforced range. This finding supports the notion that adult nonsingers may be reliable at self-assessment of their range, which was not found to differ from assessment led or reinforced by the teacher.

Despite the large adult range relative to children's range, both children and adults have been shown to "pitch" songs in the lower end of the vocal range when not given a reference pitch. Elementary students as well as elementary education majors sang low in their respective ranges given a familiar song (Moore, 1991), suggesting an individual's range does not predict the starting pitch when singers self-select a starting pitch. However, Guise (2003) tested undergraduate choir members by asking them to sing a self-selected pitch and report their voice part. They found there was a significant difference between sopranos and altos' self-selected starting pitch, and between tenors and basses' starting pitch. Thus, starting pitch selection may be influenced by experience, practice, and experience singing specific voice parts.

Starting pitch selection for familiar songs may be based on its features and possibly also the context in which it was learned. Some research findings have indicated a cultural impact on mean speaking pitch (F_0) and mean singing pitch, namely that F_0 was significantly lower than singing pitch, particularly for English monolinguals (Mang, 2002). However, mean singing pitch for songs learned from a vocal model like "Happy Birthday" were sung at a higher pitch than a choice. These are typically sung privately, suggesting an influence of models heard in daily life or in instructional settings. For certain tunes often heard in particular keys, individuals may be more likely to sing in a common key, and do so again on a retest (Halpern, 1989). Between 25% and 44% of individuals can be expected to sing a famous pop song in the key it was originally performed (Frieler et al., 2013; Levitin, 1994). However, for songs with less key center expectancy, as in ones we tend to sing along with but do not often listen to for pleasure (such as "Happy Birthday"), selection of the starting pitch may be more dependent on features specific to the song or the individual singing it.

Starting pitch selection—whether intentional or habitual or otherwise—has been theorized to vary according to the key center and opening contour of a melody (Ogawa, 1997). Additionally, pitch sequences beginning in a distant scale degree might influence one's starting note regardless of the key center. Starting low in one's vocal range is a persistent characteristic of undergraduate students, sometimes even when melodies begin on a higher note (Ogawa, 1997). Ogawa noted that in previous studies (e.g., Moore, 1991), singers did not take general melody characteristics such as contour into consideration. When junior early childhood education majors sang 13 familiar Japanese songs, the students tended to select a

lower starting pitch when the melody's starting pitch range was low. When the song had a higher beginning pitch contour, some students began low in the voice, whereas some began higher near the original given pitch. This suggests a possible relationship between knowledge of the melody and its range and selection of starting pitch.

Undergraduate nonsinger music majors may pitch songs lower in their range than do active singers (Cevasco, 2008; Killian & Buckner, 2008). Furthermore, trained singers have a wider flexibility in their vocal range (Cevasco, 2008; Killian & Buckner, 2008; Siupsinskiene & Lycke, 2011). Moreover, self-selecting a higher pitch has been shown to lead to greater initial pitch accuracy and key accuracy (McCoy, 1997). There is evidence of a clear effect for training in the music major population and in children: skilled singers were more likely to pitch songs higher in their range than untrained singers and were also more capable of performing them accurately.

The existing research focuses primarily on music majors, elementary education majors without significant singing experience, and children, with fewer data for nonmusic majors with significant choral experience. How do undergraduate choral musicians who are not music majors place familiar songs in their vocal ranges? The purpose of this study was to explore key selection among well-known songs and whether singers choose the starting pitch intentionally. The research questions were as follows: (1) How do collegiate nonmusic major adult singers place familiar songs in their range, and (2) How intentionally do these singers report selection of a starting pitch?

Method

Participants

The participants for this descriptive study were recruited from one large mixed SATB (soprano, alto, tenor, and bass) ensemble at a public university in the Northeast. None of the participants were music majors. We followed an IRB-approved protocol for implied consent documented through an online data collection tool. Singers participated remotely using their own devices, having been encouraged to find a quiet space with no distractions and to turn off mobile notifications on their devices. The sample consisted of undergraduate students ($N = 28$) who were actively participating in a choir. The mean age was 19.57 ($SD = 1.14$). All voice parts were represented, including sopranos ($n = 7$), altos ($n = 14$), tenors ($n = 5$), and basses ($n = 2$). Finally, these students reported a wide variety of majors, with aerospace engineering, psychology, and criminology being the most popular at two students each.

Song Tasks

Four songs were chosen for their familiarity, variety in range, and differences in where the starting pitch lies respective to the rest of the song and the key center. "Happy Birthday" begins on the song's lowest note, which is the fifth scale degree below the tonic ("sol") in the key. "Row Row Row Your Boat" also begins on its lowest note, but in this case, it is the first scale degree and thus also the key center. "Jingle Bells" begins on the third scale degree, and in this case, it is not the song's lowest note, which is two scale degrees lower. Last, "The Star-Spangled Banner" possesses the highest starting pitch relative to its key center on the fifth scale degree above the tonic. "The Star-Spangled Banner" possesses the largest range of all four songs (an octave and a half), whereas "Jingle Bells" has the smallest range (a fifth).

Additional nonsinging activities included a speaking pitch frequency task, or F_0. The instructions given to participants were as follows: "Now let's record you speaking backward from 10 in your normal, regular speaking voice. Speak comfortably, like you're talking to a friend across a small room." In the following tasks, we aimed to identify the participant's vocal range, and the instructions read as follows:

> Now click this link and sing UP the scale to your very HIGHEST note using any syllable. It's okay to take a breath anywhere you want, and it's okay to pause and start again—just don't stop the recording. When you've gotten to your absolute HIGHEST singable note, hold it a few seconds, then end the recording.

The instructions for lowest note were the same but replaced the word "highest" with "lowest." Like the song tasks, this range task was presented a cappella on a syllable of the participants' choosing. Following these tasks, we asked participants questions about range characteristics of the criterion songs, such as which song had the largest range, and how intentionally participants chose the starting pitch. We also asked questions about age, voice lesson and choir history durations, and voice part assignments in choir.

Procedure

Participants completed data collection online and remotely. They were asked to record themselves singing certain familiar songs and nonsong tasks, following suggested recent practices for online data collection (Sauter et al., 2020). Songs were chosen based on familiarity, and represented various pitch ranges and previous usage in the literature (e.g., Mang, 2002; McCoy, 1997). Participants performed "Happy Birthday," "Jingle Bells," "Row Row Row Your Boat," and "The Star-Spangled

Banner" in their entirety without a reference pitch in an order randomized by the presentation software. Between each recording task, participants listened to a recording of pitched noise in a piano timbre (Nichols et al., 2018) to minimize transfer of key or pitch from the previous task. Next, they recorded themselves counting from 10 down to one as a measure of speaking voice pitch (Fisher, 2014; Killian & Wayman, 2010). Finally, they recorded themselves singing a "slide" down to their lowest pitch, and then singing up to their highest pitch. At the end of the singing portion, participants completed questions about their backgrounds and singing experience. Anonymized audio files were downloaded and stored electronically by a research assistant.

Online Data Collection

Data were collected via the web browsers used by participants: Safari ($n = 17$), Chrome ($n = 24$), and Edge ($n = 2$). Participants consented via the online instrument regardless of whether they subsequently submitted completed responses. Further data were collected concerning which operating system was used by participants, including Mac ($n = 17$), Windows ($n = 11$), Android (2), iPhone ($n = 12$), and other ($n = 1$). Cooke et al. (2011) reviewed two concerns for web-based studies involving audio: experimental control and trustworthiness of responses. We acknowledge that data were not collected in sound-attenuated studios with state-of-the-art equipment for audio recording. However, analysis of pitch frequency yielded scorable data for each of the criterion vocal tasks. When evaluating pitch using Praat software (Boersma & Weenink, 2018), the only reliability concern was in participants' tendency to shorten or speak/sing short syllabus like the first syllable of "Happy Birthday," for which we corrected by using more reliable syllables described later. Second, we estimate the trustworthiness of participants' effort to be high due to recruitment from an elective class of singers.[1]

Scoring

The participants' recordings were analyzed for acoustic measurements of pitch using Praat (Boersma & Weenink, 2018) software to determine the pitches produced for the song and nonsong tasks. Measurements were taken by locating the starting pitch and isolating the area of the pitch with the flattest slope (the middle stable portion). Thus, these pitch measurements were captured manually, one-by-one and recorded in Hz and also converted to a pitch class designation (note name, e.g., F#3) by rounding up or down to the nearest pitch by using the A-440Hz standard. A second scorer evaluated 15% of the data, including nonsong tasks, and interrater reliability was deemed acceptable ($r > .85$). We transformed frequency in Hz to cent deviations from C2, a stable marker which was lower than any values yielded in the data.

For analysis, we used the second pitch for "Happy Birthday" and the third pitch for "Jingle Bells" for a more reliable evaluation of starting pitch not confounded with scooping or tendency to speak-sing the first syllable. For "Happy Birthday," measurements were taken on both syllables of "happy" and it was found that the second syllable "-py" was more consistently flat in slope (less scooping compared with the first). The same process occurred with "Jingle Bells" for the first three pitches, with "bells" found to be the most consistent (i.e., least tendency for scooping). For "Row Row Row Your Boat," an average of all the "rows" was taken. For "The Star-Spangled Banner," the pitch of the third note was used ("say") because it was a more stable indicator, and the starting pitch (dominant pitch) was estimated from that note, which was the tonic. No attempt was made to determine key stability nor singing accuracy in terms of intonation across the duration of the excerpt.

In scoring the nonsong tasks, we followed the same strategy of exploration and analysis. To document the F_0 for speaking pitch, participants counted backward from ten to one. We examined the last two numbers, two and one, to allow for stabilization of the speaking pitch across the initial span of counting. However, the frequency graph for "one" indicated consistent scooping, making it unsuitable for measuring mean frequency during the pitch. Instead, the number "two" was a more stable measure, and the speaking voice pitch measurement reported for this task is taken from the number "two" for each participant.

For the lowest and highest notes in the participant's vocal range, the measurement was taken on the last terminal pitch achieved measured in Hz. Because the participants were asked to hold the final note for a few seconds, the last note had a flat, stable slope and could be measured accurately. In particular for the highest note task, it was noted by the first rater that all but one of the male participants accessed their falsetto voice when singing to their highest note.

Exploratory Analyses and Results

To best represent the central tendencies using a linear scale, and to minimize the impact of outliers, we report the median pitch for each task and the median absolute deviation. Because Hz is a continuous and logarithmic scale for frequency and pitch class is discrete, we converted frequency measurements to cents, a 100-equal-unit measure of half-step intervals we chose to number from C2, a pitch lower than any found in the analysis.

Table 1. Median Pitch Class for F0, Range, and Starting Pitch Selection (Median Absolute Deviation in Cents).

	n	Speaking (F0)	Lowest pitch	Highest pitch	HBD	RRR	SSB	JB
Soprano	7	G#3 (147)	E3 (127)	C6 (204)	B3 (14)	B3 (70)	D#4 (107)	E4 (205)
Alto	14	G3 (125)	D#3 (213)	A5 (164)	B3 (121)	C4 (121)	D#4 (167)	E4 (106)
Tenor/bass	7	B2 (99)	G2 (56)	D#5 (353)	C3 (65)	C3 (0)	F3 (275)	F#3 (113)

Note. HBD = Happy Birthday; RRR = Row Row Row Your Boat; SSB = Star-Spangled Banner; JB = Jingle Bells.

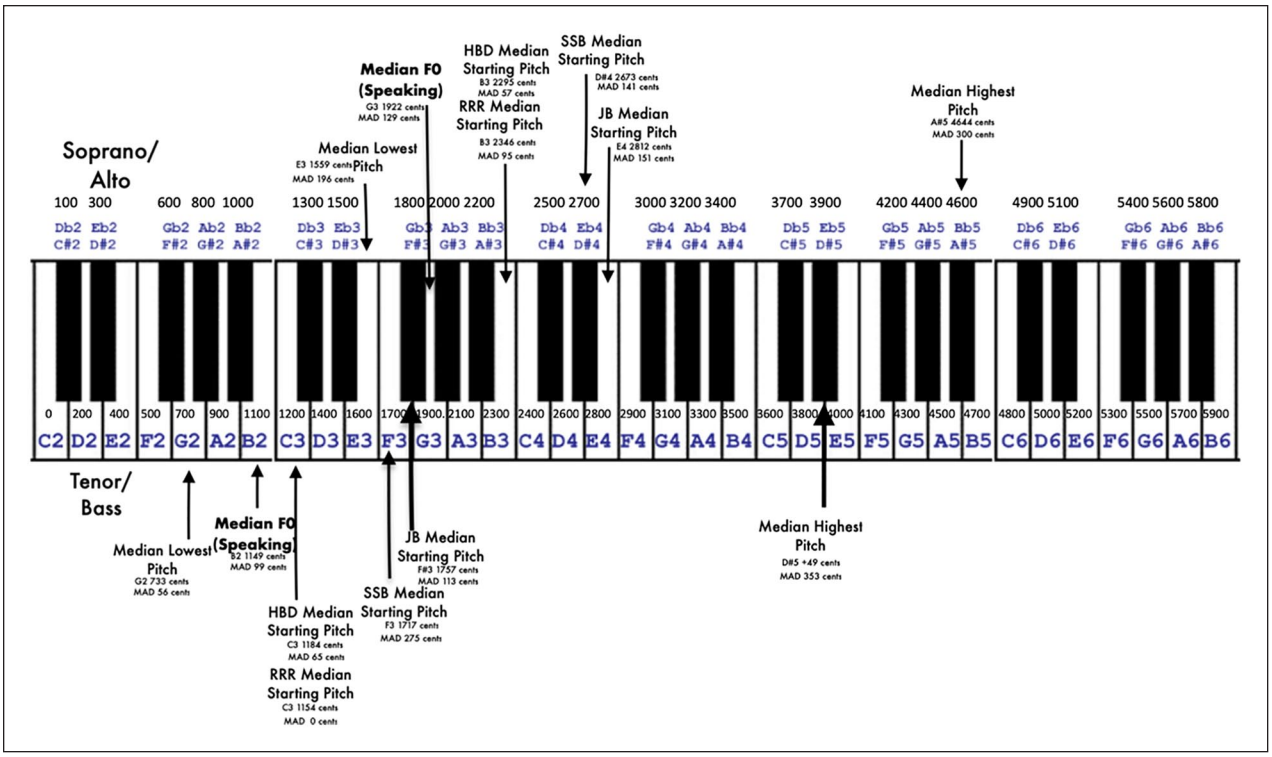

Figure 1. Median values in cent deviation with median absolute deviation (MAD) for speaking (F_0), lowest pitch, highest pitch, and starting pitch for each song.
Note. Treble clef singer data are above the keyboard and bass clef singers are below. HBD = Happy Birthday; RRR = Row Row Row Your Boat; SSB = Star-Spangled Banner; JB = Jingle Bells.

Next, we calculated pitch class data by manually assigning pitch class values to the cent deviation scores. First, we present in Table 1 the pitch class data by voice type in accordance with previous reporting practices in music education (Killian & Buckner, 2008; Killian & Wayman, 2010). To better represent the relationships among tasks, we report pitches as a number of cents deviating from C2, separately for treble and bass clef voice types. We chose to express these values in relation to the standard keyboard in Figure 1 for an ecological expression of voice usage and to facilitate comparison with previous research in music education.

A preliminary analysis of median values indicated the median pitch of bass clef voice types (tenor and bass) was not precisely 1200 cents (one octave) lower than the median pitch of treble clef voice types (soprano and alto) among the various tasks. We continued by excluding participants indicating bass clef voices, as they could not be combined with treble clef voices for analysis theorized as displaced by one octave. Thus, to avoid a confound of range differences between treble and bass clef voice types and for comparison with previous research using adult treble voices, we chose to indicate only the range selection of singers who indicated treble clef voice types (n = 21). Figure 2 represents the placement of each individual treble voice in all four song tasks.

After the song and nonsong tasks, all participants were asked how carefully they chose the starting pitches of the songs they performed, with the option to answer *Not really/I don't know* (n = 12), *Somewhat* (n = 12), or *Yes* (n = 4). Similarly, participants were asked if they intentionally chose a range for their specific voice

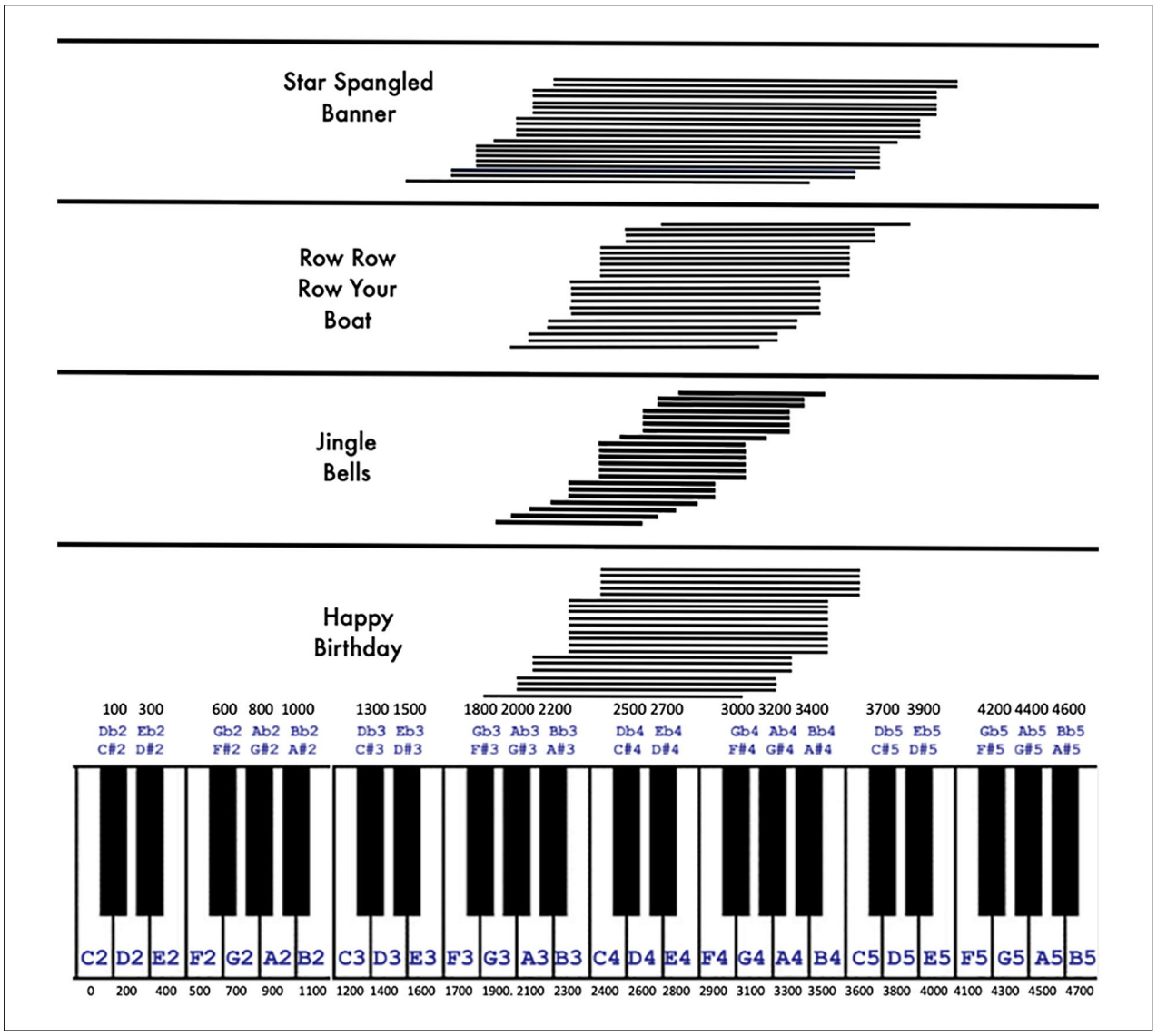

Figure 2. Placement in the voice range by soprano and alto participants on the four song tasks ($n = 21$).

and responded, *Not really/I don't know* ($n = 5$), *Somewhat* ($n = 14$), or *Yes* ($n = 9$). Participants indicated the song they believed had the largest range: "The Star-Spangled Banner" ($n = 27$) and "Row Row Row Your Boat" ($n = 1$). Finally, participants indicated the song they believed had the smallest range: "Jingle Bells" ($n = 16$), "Row Row Row Your Boat" ($n = 8$), or "Happy Birthday" ($n = 4$).

These collegiate nonmusic major singers indicated whether they had ($n = 19$) or had not ($n = 9$) previously taken voice lessons. Of participants indicating a history of voice lessons, responses ranged from one to nine ($M = 3.89$ years, $SD = 2.37$). We asked how many years participants sang in a choir including the present year, with responses ranging from one to 16, indicating a history of singing practice and experience ($M = 8.22$ years, $SD = 4.39$). To explore the association between voice lesson history and other background questions, a series of likelihood ratio tests were performed. Results indicated a history of voice lessons was not related to participants' responses to the item, "I chose starting pitches carefully," $\chi^2(2, N = 28) = .869$, $p > .05$. Voice lesson history was also not related to responses to the statement, "I did not think much about the key," $\chi^2(1, N = 26) = 0.110$, $p > .05$. Results indicated a history of voice lessons was associated with a positive response to "I chose an intentional range for my voice," $\chi^2(2, N = 28) = 6.222$, $p = .045$.

Discussion

The purpose of this study was to explore the starting pitch selection of undergraduate singers when asked to sing

four familiar songs. The participants represented an important part of the music school community: individuals with rich singing experience but who were not majoring in music. They can be expected to vary in terms of their formal music knowledge, including notation reading ability and music theory or history. When asked which of the four song tasks represented the largest ranges of pitches ("The Star-Spangled Banner"), all but one participant answered accurately. A majority of participants ($n = 23$) reported they intentionally or somewhat intentionally chose song ranges to fit their voices. Thus, participants indicated an awareness of range features, even if not all carefully chose what key to perform each song. Importantly, a history of voice lessons was associated with a positive response to choosing an intentional range, suggesting that formal music training experiences may yield a more nuanced approach to considerations in performing a song.

Over half the participants self-reported having intentionally chosen the starting pitch or range for the songs they performed. We suggest participants were more likely to respond affirmatively based on their singing background, which differs from previous findings of collegiate nonvocalist elementary education majors (Killian & Buckner, 2008; Kuhn et al., 1979; Ogawa, 1997). Furthermore, the participants in the present study may have responded based on an expectation that they should consider range or starting pitch to be important for singing. An alternate design might have parsed these items differently; we chose to ask these questions following the singing tasks rather than before so as to avoid priming participants to think about range or pitch more intentionally than would be expected for spontaneously sung vocal music in a solo setting.

Treble voice participants began each song within a small distribution. The median absolute deviation values for starting pitches ranged 100 to 200 cents (one to two half-steps) within each song task. With the exception of "The Star-Spangled Banner," bass voice participants also indicated median absolute deviation values between 100 and 200 cents for song tasks. This finding is presumed to not be influenced by participation in a specific choral ensemble because these songs are not representative of the standard choral repertoire. Therefore, we suggest these findings were not derived from unseen effects of the specific choral ensemble from which the participants were recruited.

Previous research indicates variability in the speaking voice pitch of individuals (Siupsinskiene & Lycke, 2011), and we expected greater variability in the starting pitches chosen. One possible explanation is that participants with significant singing experience may possess the ability to place song ranges within a comfortable vocal range with acuity. Participants were generally able to complete the songs successfully within their range after beginning, including accurate pitches and rhythms. Thus, these collegiate singers possessed the ability to choose an appropriate starting pitch which resulted in a range that fit their voices, even in the absence of self-reported intentional key selection.

The median starting pitch for treble voices ranged from B3 ("Happy Birthday," "Row Row Row Your Boat") to D#4 ("The Star-Spangled Banner") to E4 ("Jingle Bells"). Nonmusic majors who were presumed not to have significant choral singing experience in a previous study demonstrated a mean starting pitch of B-flat3 on the songs "Happy Birthday" and "Twinkle, Twinkle" (Killian & Buckner, 2008) and A-flat3 on the song "Row Row Row Your Boat" (Killian, 1996). As in those previous studies, participants of the current study also began low in the singing range for two of the songs, "Happy Birthday" and "Row Row Row Your Boat." However, for the songs "The Star-Spangled Banner" and "Jingle Bells," the median starting pitches were higher (D#4 and E4, respectively). For these two songs, the median pitch can be said to be in the middle range between C4 and C5, suggesting starting pitch selection is dependent on the song selected. We interpret these data to suggest that participants placed songs in the middle portion of the vocal range, making them different than previously reported nonsingers on similar tasks and including previous research using different tasks (Moore, 1991). The results reinforce earlier findings that active choral music participation may result in a greater use of the vocal range (Siupsinskiene & Lycke, 2011).

In regard to vocal ranges, singers employed ranges closer to the bottom of the vocal range than the top, but not as low as nonsinger collegiate populations reported in previous research (Killian, 1996; Killian & Buckner, 2008). Previous nonmusic major participants generally had less choral experience and were selected from academic music courses, not an ensemble (Cevasco, 2008). They had an average of 5 years of choral experience compared with our participants, who had an average of 8 years and were all in choir at the time of participation. Accordingly, the present nonmusic major singers had wider vocal ranges: compared with a span from G3 to B4, treble clef singers in the present study spanned D#3 to C6.

The data indicate the potential that these college musicians chose a similar starting pitch for "Happy Birthday" compared with music majors in a previous study, which was higher than that selected by nonsinger students in the same study (Killian & Buckner, 2008). However, for "Jingle Bells," participants in the present study demonstrated a median starting pitch of E4 compared with lower starting pitches in music majors and nonsingers in the previous study (between C and C3 in Killian & Buckner,

2008). For lowest pitch in the singing range, Geringer et al. (1980) also designed E3 to F3 as a central value for lowest terminal pitch in nonmusic majors, but a somewhat higher upper limit: D5 to F#5 as upper terminal pitch. When compared with nonsingers, college nonmusic major singers in the present study may have demonstrated higher voice ranges, though it is unclear how they may compare with music major singers on these specific familiar songs or among the choral repertoire.

Vocalists may choose to perform songs in any key, especially when performing alone: songs do not have to be performed in the key they are published, nor do they have to be sung in commonly heard keys. In fact, when individuals are tasked with performing a song without a reference pitch, they may begin spontaneously and sometimes without actively choosing a specific starting pitch. These college musicians indicated greater discrimination between songs when choosing a starting pitch than previously reported, which may be based on the specific song stimuli chosen for this study. Killian and Buckner (2008) suggested significant key selection variability within groups of children, music majors, and nonmusic majors without specified choral singing experience. In comparison, results of the current study indicated much less variability in college choral musicians who are not music majors.

We chose to use cents as measurements deviating from C2 as a measure of equal demarcations between half-steps to transform the logarithmic scale of Hz to a linear scale. Killian and Buckner (2008) addressed the problem of nonlinear scale by assigning integers for half-step increments (positive or negative) from C4 (e.g., +1, +2, −1 for scoring), and Killian (1996) used the same approach but with E4 as the baseline reference pitch. Each of these approaches seemed to provide a productive strategy for analyses, and our method was based on a decision to report medians and median absolute deviations using unsigned values. We suggest this is also a valid strategy, including the choice to indicate performance in terms of deviation from a note low in range (C2, lower than the typical vocal range). We suggest the use of median absolute deviation for cent data as an indicator of variability in research based on data transformed from a logarithmic scale such as Hz. Furthermore, the use of medians for reporting central tendency has been useful for describing small sample sizes in descriptive research like the present study and in previous research (Killian & Buckner, 2008).

We use the comparisons with previous research to suggest that remote data collection was a reliable method for evaluating vocal pitch range in a sample of college musicians because we found the expected larger vocal range in active singers. Our remote data collection relied on sampling from one institution; it did not represent a crowdsourcing approach which is sometimes used for remote data collection. We would likely have sampled from intact, same, or similar courses if conducting in-person data collection. We presume these participants are as representative of the population as lab-based participants because we recruited from a source class at one institution (c.f. Woods et al., 2015). However, this type of data collection may not be reliable in populations of children or nonsinger adults, where a teacher or research assistant's support may be necessary to help novice singers explore their ranges.

Participants judged song ranges based on knowledge of the songs and experience singing them and could generally name "The Star-Spangled Banner" as the largest range and "Jingle Bells" as the smallest range. Starting pitch selection on unfamiliar, newly learned songs was not explored in the present study, and we did not attempt to ascertain participants' level of familiarity with each song during recruitment. Students who chose to participate in the study were able to perform each of the song tasks; however, students from the dominant culture may have been more likely to participate and would be more likely to have considered these four songs familiar.

Limitations and Future Research

The choice to avoid comparisons of the bass clef voices to the treble clef voices is important because the analysis indicated differences of median starting pitch varied by less than an octave between the two groups. For example, the median starting pitch of bass clef voices was a minor seventh lower than treble clef voices for "Jingle Bells" and "The Star-Spangled Banner," and a major seventh lower for "Happy Birthday" and "Row Row Row Your Boat." For speaking voice medians, bass clef voices were a minor sixth lower than treble clef voices. Therefore, the concept of octave displacement for bass clef voices in previous research may present an imprecise expression of bass clef singer ranges in relation to treble clef singer ranges. In the present data, the difference is less than octave, suggesting bass clef voices cannot be compared directly with treble clef voices as same-but-lower. These results are limited to findings in college singers who are not music majors and presents a challenge for future researchers seeking to incorporate the bass clef voice population.

Future work can be used to develop remote testing measures with high reliability for assessing the highest and lowest terminal pitches in a singer's range. Geringer et al. (1980) suggested reliable self-assessment of range in populations similar to but different from the population in the present study. However, the possibility remains the present data reflect a self-reporting confound in which singers may not have actually reached the lower

or upper limits of the voice range without the presence or guidance of a teacher or of researchers. The lower and upper range limits in the present study should be interpreted with caution.

Implications for Music Education

The questionnaire items for song range consideration and whether participants chose starting pitches carefully were introduced after the participants completed the singing tasks; we did this intentionally to reduce priming the students to consider range differently than they would for solo song tasks at their discretion. Thus, they were not primed to consider range or starting pitch any more than they typically might in an educational or performance setting. However, it is possible that choral singers recruited from a campus choir might feel expected to have made such considerations, and possibly even more so for singers with a history of voice lessons. Therefore, we suggest the responses to these items may not be reliably interpreted: students may have been influenced by a personal or perceived expectation that they should have considered the range of the songs or their own voice range based on their music experience or status as active singers.

Elements of music theory are variably taught in ensemble courses, including those that may include music majors and nonmusic majors. If group performance is the primary goal, the curriculum may not generally afford instruction to students for vocalizing and performing independently of the ensemble. Additionally, ensemble directors in choral settings often determine the voice part to which the singer will be assigned, and knowledge of one's own voice or range may not be required of the singer. The ensemble curriculum may not include music components that would lead students to develop skills in "pitching" a song for individual performance.

These results may suggest the choral music curriculum at elementary, secondary, and postsecondary levels may be adjusted so singers learn about their individual voices, including identifying their voice type, preferred voice part, and choosing a key when singing familiar or unfamiliar songs. For singers at those age levels who do not take private lessons, the choir curriculum may represent the only music content in which the student engages. Developing individual musicianship skills such as range knowledge and key considerations would allow students to continue performing successfully, whether alone or in an ensemble, after their participation during school ensemble years.

Compared with children, adult singers have been shown to have increased vocal capabilities such as pitch range and upper note range limits (Siupsinskiene & Lycke, 2011). The median pitch for lowest terminal pitch of the voice range was E3 for treble clef voices, which matches that published previously for nonmusic major nonsingers (Kuhn et al., 1979). However, the median pitch for highest note in the present study of the range was B-flat5, a fifth higher than for nonmusic major nonsingers in the previous study. Our finding supports the previous authors' assertion that singing experience can and should include range expansion as part of the curriculum, or that singing experience may yield effects for singer range expansion whether directly taught or otherwise.

In conclusion, the main findings are that collegiate singers employed starting pitches that were remarkably close to one another and may have chosen pitches in the middle—rather than lower—portion of the singing range. The results indicated some participants did consider the starting pitch or range of the song stimuli, which might be expected to be a part of the vocal music curriculum. Furthermore, the starting pitch selection of these participants cannot be assumed to be similar to populations of children, music majors, or nonmusic major nonsingers. Though frequently enrolled in collegiate ensembles, both conductor and student-led, the population of collegiate nonmusic majors with significant singing experience such as the participants in the present study have not been a focus of previous research in singing. However, their presence in collegiate ensembles reflects the transition from the K–12 music experience to potentially life-long music making across the life span.

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.

ORCID iD

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

Note

1. Previous research indicated successful deployment of an online tool for conducting spoken word research online (Slote & Strand, 2016), though reporting a much higher ratio of Windows to Mac platforms. Those authors suggested online data collection may help avoid experimental bias or effects of participant expectation. Further to bias effects, which should be important in vocal performance research, test administration factors including the delivery of instructions and procedures can be made identical for every participant. Other psycholinguistic research has also reported high correlation between data collected online

using Amazon Mechanical Turk and traditional methods (Schnoebelen & Kuperman, 2010).

References

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer* (Version 6.0.43) [Computer program]. www.praat.org/

Cevasco, A. M. (2008). Preferred vocal range of young and older adults: Implications for music therapy majors' clinical training experience. *Music Therapy Perspectives*, *26*(1), 4–12. http://doi.org/10.1093/mtp/26.1.4

Cooke, M., Barker, J., Lecumberri, M. L. C., & Wasilewski, K. (2011). Crowdsourcing for word recognition in noise. *Interspeech*, *2011*, 3049–3052. www.isca-speech.org/archive/archive_papers/interspeech_2011/i11_3049.pdf

Demorest, S. M., Kelley, J., & Pfordresher, P. Q. (2017). Singing ability, musical self-concept, and future music participation. *Journal of Research in Music Education*, *64*(4), 405–420. http://doi.org/10.1177/0022429416680096

Fisher, R. A. (2014). The impacts of the voice change, grade level, and experience on the singing self-efficacy of emerging adolescent males. *Journal of Research in Music Education*, *62*(3), 277–290. http://doi.org/10.1177/0022429414544748

Frieler, K., Fischinger, T., Schlemmer, K., Lothwesen, K., Jakubowski, K., & Müllensiefen, D. (2013). Absolute memory for pitch: A comparative replication of Levitin's 1994 study in six European labs. *Musicae Scientiae*, *17*(3), 334–349. https://doi.org/10.1177/1029864913493802

Geringer, J., Nelson, J., & Kostka, M. (1980). Differential assessment of child and adult singing ranges. *Contributions to Music Education*, *8*, 39–46. www.jstor.org/stable/24127432

Guise, P. E. (2003). The relationship between voice part and self-selected pitch. *Phenomenon of Singing*, *4*, 94–98. https://journals.library.mun.ca/ojs/index.php/singing/article/view/565

Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, *17*(5), 572–581. https://doi.org/10.3758/BF03197080

Killian, J. N. (1996). Definitions of "knowing": Comparison of verbal report versus performance of children's songs. *Journal of Research in Music Education*, *44*(3), 215–228. https://doi.org/10.2307/3345595

Killian, J. N., & Buckner, J. J. (2008). Comparison of starting pitch preferences among fourth graders, undergraduate music majors and elementary education majors. *Music Education Research International*, *2*, 11–20. http://cmer.arts.usf.edu/content/articlefiles/962-MERI02pp11-20.pdf

Killian, J. N., & Wayman, J. B. (2010). A descriptive study of vocal maturation among male adolescent vocalists and instrumentalists. *Journal of Research in Music Education*, *58*(1), 5–19. http://doi.org/10.1177/0022429409359941

Kuhn, T., Wachhaus, G., Moore, R., & Pantle, J. (1979). Undergraduate nonmusic major vocal ranges. *Journal of Research in Music Education*, *27*(2), 68–75. http://doi.org/10.2307/3344893

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Attention, Perception, & Psychophysics*, *56*(4), 414–423. https://doi.org/10.3758/BF03206733

Mang, E. (2002). An investigation of vocal pitch behaviors of Hong Kong children. *Bulletin of the Council for Research in Music Education*, *153/154*, 128–134. www.jstor.org/stable/40319152

McCoy, C. W. (1997). Factors relating to pitch-matching skills of elementary education majors. *Journal of Research in Music Education*, *45*(3), 356–366. http://doi.org/10.2307/3345531

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. www.jstor.org/stable/40318417

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

Nichols, B. E., Wöllner, C., & Halpern, A. R. (2018). Score one for jazz: Working memory in classical and jazz musicians. *Psychomusicology: Music, Mind, & Brain*, *28*(2), 101–107. https://doi.org/10.1037/pmu0000211

Ogawa, Y. (1997). Starting pitch and unified key process in adult group singing. *Bulletin of the Council for Research in Music Education*, *133*, 100–104. www.jstor.org/stable/40318847

Sauter, M., Draschkow, D., & Mack, W. (2020). Building, hosting and recruiting: A brief introduction to running behavioral experiments online. *Brain Sciences*, *10*(4), 251–261. http://doi.org/10.3390/brainsci10040251

Schnoebelen, T., & Kuperman, V. (2010). Using Amazon Mechanical Turk for linguistic research. *Psihologija*, *43*(4), 441–464. http://doi.org/10.2298/PSI1004441S

Siupsinskiene, N., & Lycke, H. (2011). Effects of vocal training on singing and speaking voice characteristics in vocally healthy adults and children based on choral and nonchoral data. *Journal of Voice*, *25*(4), 177–189. http://doi.org/10.1016/j.jvoice.2010.03.010

Slote, J., & Strand, J. F. (2016). Conducting spoken word recognition research online: Validation and a new timing method. *Behavior Research Methods*, *48*(2), 553–566. https://doi.org/10.3758/s13428-015-0599-7

Wolf, D. L. (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. www.jstor.org/stable/40311596

Woods, A. T., Velasco, C., Levitan, C. A, Wan, X., & Spence, C. (2015). Conducting perception research over the Internet: A tutorial review. *PeerJ*, *3*, e1058. http://doi.org/10.7717/peerj.1058