Perceptual Measures of Boychoir Voices During the Phases of Pubertal Voice Mutation

[*,†]Barbara Weinrich, [*,†]Susan Baker Brehm, [‡]Wendy LeBorgne, [§,1]Christopher Eanes, [*,2]Stephanie Zacharias, [*]Janet Beckmeyer, [‖]Michael Hughes, and [¶]Alessandro de Alarcon, *‡§¶*Cincinnati, and* †‖*Oxford, Ohio*

Summary: Background/Objectives. Vocal changes in the male singing voice associated with puberty are variable and often unpredictable resulting in challenges for the singer and the choral director. Limited knowledge regarding the physiologic changes in the vocal mechanism as they correlate to perceptual variations observed in the male adolescent singer exists in the literature. The purpose of this study was to examine pitch breaks and perceptual characteristics of vocal quality during singing tasks for boys in various stages of the male changing voice.

Study Design. Prospective Study.

Methods. Twenty-eight boys were initially evaluated at Cooksey Stage 0 (Pubertal Unchanged; n = 15) or Cooksey Stage 1 (Mid-Voice; n = 13). Range of age was 8−13 years old. Participants performed vocal slide intervals (1-3-1, 1-5-1, 1-8-1) with discrete starting frequencies on G3, C4, F4, and A4 and sang the “Star-Spangled Banner” in the key of Ab. Pitch breaks and perceptual qualities were evaluated on the recorded tasks by expert raters. Seven boys were evaluated again when they progressed to Cooksey Stage 4 (Baritone) performing the same singing tasks.

Results. For the participants evaluated at Cooksey Stage 0/1, pitch breaks were observed more in the higher frequencies and increased interval spacing regardless of starting frequency. Participants at Cooksey Stage 0 had more pitch breaks than Stage 1. At Cooksey Stage 4, an increase in the number of pitch breaks was observed in comparison to their tasks performed at Stage 0/1 and the perceptual quality of breathiness was significantly greater.

Conclusions. Pitch breaks are a characteristic perceptual change that indicates a young man may be transitioning through puberty. Findings from the present study demonstrate that in addition to perceived pitch breaks, breathiness was noted to significantly increase as the male progressed through puberty. Breathiness was noted to be more significant than vocal timbre and overall vocal quality. This research provides acoustic evidence to enhance the perceptual characteristics of voice change for those who teach and train male voices through puberty.

Key Words: Boychoir−Singers−Pitch breaks−Perceptual measures−Voice change−Puberty.

INTRODUCTION

Choral singing activities provide a primary vocal outlet for young boys to explore the art of singing. The understanding of both the physiologic and perceptual vocal changes through puberty are necessary for optimizing healthy and safe vocal training practices during this tenuous time of development. However, there is limited data on physiologic and acoustic changes in the male voice during puberty, supporting the need for further investigation into this topic. The few studies that document pubescent vocal development in this population do not include current relevant measures of voice and vocal production which can be utilized. Specifically, the relationship between perception and acoustic measures has not been documented in the literature. Additionally, the manner in which data has been collected and evaluated varies by clinical site and has not been consistently applied and documented in previous studies. Although there is a significant research base related to physiologic changes in both the physical body and vocal mechanism, few studies have examined perceptual measures associated with corresponding voice change. The limited data from a vocal training/pedagogy standpoint is primarily anecdotal interpretations and personal life experiences in working with and training young male voices.

Historically, the pedagogical training of the male singing voice through the period of puberty is often the result of singing voice teachers' experience and understanding of what they perceive in the absence of accurate physiologic understanding of the laryngeal mechanism.[1−7] The master-apprentice relationship of vocal training found in the pedagogical literature holds true for the adolescent male voice as well. Traditionally, some singing teachers and choral directors were hesitant to train the male voice through puberty due to concern of inducing vocal damage, and perhaps due to lack of formal knowledge of the structure and function of the voice as it changes. The unpredictable nature and quality of the male voice during this timeframe make it difficult

Accepted for publication April 13, 2020.
Conflict of Interest: None.

From the *Division of Speech-Language Pathology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; †Department of Speech Pathology & Audiology, Miami University, Oxford, Ohio; ‡Professional Voice Center of Greater Cincinnati, Cincinnati, Ohio; §Cincinnati Boychoir, Cincinnati, Ohio; ‖Department of Statistics, Miami University, Oxford, Ohio; and the ¶Division of Pediatric Otolaryngology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio.

[1]Cathedral Choral Society, Washington National Cathedral, 3101 Wisconsin Ave NW, Washington, DC 20016, USA.

[2]Head and Neck Regenerative Medicine Laboratory, Mayo Clinic Arizona, Division of Pediatric Otolaryngology, Phoenix Children's Hospital, Phoenix, AZ, USA.

Address correspondence and reprint requests to Barbara Weinrich, Department of Speech Pathology & Audiology, Miami University, 1605 Gage Drive, Middletown, OH 45042 USA E-mail addresses: weinribd@miamioh.edu, Barbara.Weinrich@cchmc.org

Journal of Voice, Vol. 36, No. 1, pp. 142.e1−142.e8
0892-1997

https://doi.org/10.1016/j.jvoice.2020.04.002

for professional choirs to adequately maintain consistency within the choir. Beginning with the tradition of Boychoirs in Leipzig, from 1727 to 1749, the average age of voice change was assessed to occur around 18 years of age.[8] The average age of voice change in males has lowered throughout the centuries and during the 1960's and 1970's in Europe was estimated to occur between 14.2 and 15.5 years. The downward trend has continued with the most recent reporting (1994−2003) estimating the average age of male puberty at 14 years.[9,10] Juul et al retrospectively examined average age of voice change in 463 boys within the Copenhagen Royal Boychoir using a body mass index comparison. Their findings indicated that a higher body mass index prepuberty correlated to an earlier age of voice change. It is suspected that the age of onset of puberty in the United States is less than 14 years[11] with WebMD (2012) reporting ages as low as nine to ten years for U.S. males. No studies to date have looked at voice change within the American tradition of boychoir singing.

To quantify physical and vocal changes, both the scientific and artistic community have established rating scales to document the progression of puberty.[4,12−14] In the musical literature, John Cooksey, a choral director, established a set of guidelines to assess the male singing voice based on acoustic voice features (eg, timbre, range, pitch breaks) through puberty. These are known as the Cooksey Stages (Appendix A; 5 Stages of Male Changing Voice following Unchanged Voice). Cooksey Stages are based primarily upon mean vocal range and tessitura (comfortable range with best vocal texture or timbre.) Tanner addressed the physical changes by describing five stages of male pubertal development that categorized the anatomical and physiologic changes associated with puberty from Stage 1 (beginning of hormonal changes with no physical markers) through Stage 5 (completion of puberty). One study has specifically compared the Tanner Stages with the Cooksey Stages in 26 nonsinger boys between the ages of 13 and 14 years. Each boy was assessed 5 times in the course of 12 months.[15] Specific assessments included height, pubertal stage via Tanner staging, testis volume, salivary testosterone profile, mean speaking frequency, mean singing frequency, speaking and singing ranges, laryngoscopy, and electroglottography. The results of this study indicated the largest height and weight changes occurred during Cooksey Stage 4 (New Voice-Baritone) and Cooksey Stage 5 (Emerging Adult Voice). Speaking and singing frequency changes were large and most significant between Tanner Stage 3 and Tanner Stage 4, and were more gradual on Cooksey's scale, occurring between Cooksey Stage 3 (Mid-Voice IIA), Cooksey Stage 4, and Cooksey Stage 5. It is noteworthy that the most dramatic vocal change in the male voice tends to occur between Tanner Stage 3 and Tanner Stage 4, indicating that dropping of fundamental frequency occurs late in puberty. Acoustic changes (timbre, frequency range) may provide a means of measurable voice change earlier in puberty before the perceived "pitch break."[1−5] Because the Cooksey Staging system is most commonly used by singing teachers and choral directors, this study provides some validity to using it as a tool to monitor vocal change through puberty.

The potential perceived invasive nature of physical examination to establish pubertal changes in boys has resulted in several self-assessment rating scales.[16−18] Results which specifically correlated the Tanner Stages with the self-rating Pubertal Development Scale[18] for 323 10- to 12-year-olds revealed a strong correlation between the two when it was supported by education of individuals on puberty changes. Body image distortion is common in this age group.[16]

There are well established studies which have documented acoustic change in male and female voices.[6,19−26] Measurements used to document change have included voice range profile, long-term average spectra, fundamental frequency (F0) over time, and speech-range profile. Findings relevant to the present study indicate that boys were able to increase their maximum vocal intensity up to the age of 10 years and then a decrease was noted in maximum decibels.[27] Also, there was a noted decrease in fundamental speaking frequency and frequency range between ages 8 and 9 years. This study indicates that vocal premutation may occur as early as ages 8−9 years in the United States with voice mutation occurring between ages 10 and 11 years.

Establishment of vocal changes in the male voice during puberty is of high interest to choral directors, particularly those working with boys in a range of ages. One study, conducted within the Copenhagen Royal Boychoir, had weekly voice assessments completed by the conductor to determine if there were any unintentional changes in falsetto or vocal tone/timbre. The date of the vocal change was established for each singer and registered by the choral assistant. These students were then taken immediately for a fiberoptic laryngeal examination to ensure that there was no overt laryngeal pathology causing the voice change. This is the only documented study to report a weekly perceived quality assessment for vocal change. However, no supporting acoustic data accompanied the reported results.[28]

There is a paucity of literature related to the perceptual aspect of vocal change in the male singing voice. Knowledge of perceptual changes may guide and inform choral directors regarding expected changes in quality and pitch breaks during development. The purpose of the current, longitudinal study, which was conducted for a period of 33 months, was to expand the documentation of the soprano (unchanged/prepubertal) male singing voice regarding pitch break and more in-depth perceptual qualities. Additionally, this study examined pitch break and perceptual quality differences in singing voices categorized as Cooksey Stage 0/1 versus Cooksey Stage 4. Information gained through perceptual correlates to physiologic changes may assist in guiding choral directors through the tumultuous vocal fluctuations associated with male voice through puberty.

MATERIALS AND METHODS

This prospective study design was reviewed and approved by the Cincinnati Children's Hospital Review Board prior to collecting data (CCHMC IRB# 2014-5211).

Participants

All participants were recruited from the Cincinnati Boychoir, which is an elite choir for young males between the ages of 7 and 17 years. Participants were classified as having a soprano voice and identified by the artistic director of the Boychoir as having a vocal range in either Cooksey Prepubertal Unchanged Voice (Stage 0) or Cooksey Mid-Voice I (Stage 1). Relevant vocal/medical history was obtained using a questionnaire at the initial visit. Participants were enrolled and evaluated, if they did not have cognitive or learning disabilities that significantly interfered with an understanding of the evaluation tasks (determined by the evaluating speech-language pathologist) and did not have any laryngeal or airway pathologies that influenced vocal quality at the time of data collection. Written informed consent was obtained from the participant's parent or guardian, and verbal assent was obtained from the participant. Participants returned for evaluation when the artistic director of the Boychoir noted a change in their vocal range that impacted their singing category (eg, soprano to alto) and their Cooksey Stage. During the 33-month time period, there were 12 boys who returned for a re-evaluation at various Cooksey Stages.

Procedures

The acoustic measures were collected in a sound-treated booth at the Center for Pediatric Voice Disorders, Cincinnati Children's Hospital Medical Center (CCHMC). These measures were obtained using the multidimensional voice program within the Computerized Speech Lab (CSL; PENTAX Medical, New Jersey). All participants stood, and voice samples were collected using a dynamic microphone (Shure SM48), with a microphone to mouth distance of 30 cm during the recording. To assess pitch breaks during vocal slides, participants sang /a/ for four or five discrete frequencies using three vocal slide intervals (1-3-1, 1-5-1, 1-8-1) for each frequency. The four frequencies used by sopranos included the notes G3 (196Hz), C4 (261.63Hz), F4 (349.23Hz), and A4 (440Hz); whereas, the five frequencies used by baritones included the notes D3 (146.83), G3 (196Hz), C4 (261.63Hz), F4 (349.23Hz), and A4 (440Hz). In addition to vocal slides, participants sang the "Star-Spangled Banner" (SSB) in the key of Ab for perceptual ratings of pitch break, overall breathiness, overall timbre, voice quality, and overall vocal quality.

Three expert judges, with greater than 10 years' experience in training adolescent singing voices, rated the singing voice samples. Each note sung in a vocal slide was rated as "yes" or "no" for a pitch break. The SSB was rated "yes" or "no" for any pitch break in the sample. For this song, the remaining perceptual measures (breathiness, timbre, voice quality, overall vocal quality) were rated with one tic mark on a 14.5 cm visual analog scale per measure. Verbal descriptors were provided for the 0 and 14.5 range as follows: overall breathiness (none—severe); overall timbre (dark—bright); voice quality (simple/light—hooty/round); overall vocal quality (poor—exceptional).

Data analyses

Data analyses were conducted using SAS software. Copyright 2002−2012 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, North Carolina.

Interrater agreement between the three raters on binary pitch break assessments was evaluated using Gwet's AC1 first-order agreement coefficient.[29] This measure has advantages over the traditional Kappa coefficient as it adjusts the overall agreement probability for chance agreement (occurs when raters agree on a rating due to one or both raters giving a random rating). For numeric endpoints (eg, breathiness), the appropriate Shrout-Fleiss intraclass correlation was used.[30] All available participant data (across all Cooksey Stages) were used in these assessments.

All subsequent comparative inferential analyses incorporated all three rater's data, and the rater effect was modeled to account for their variability in assessment response. Aggregate descriptive statistics were performed for pitch break assessments by Cooksey Stage/pitch level/vocal slide interval combinations, as well as pitch break assessments and perceptual measures in the SSB by Cooksey Stage.

For all soprano participants evaluated during Cooksey Stage 0 (Unchanged) or Stage 1 (Mid-Voice I) during their first visit, mixed effects logistic models were fit to compare the probability of pitch break between different vocal slide intervals. Separate models were fit based upon Cooksey Stage (0,1) and starting pitch for vocal slide interval combination. Dunnett-adjusted post-hoc comparisons of intervals were performed using 1-3-1 as the baseline interval. Participant and rater effects were included as random effects, and all tests were performed at the 0.05 level of significance.[31] Mixed effects models were used to compare the SSB voice perceptual assessments (Pitch Breaks, Breathiness, Timbre, Voice Quality, Overall Vocal Quality) between Cooksey Stage 0 versus Stage 1 at the first visit.

There were limited data for longitudinal comparison of voice change. Only 7 of the 12 participants who were re-evaluated achieved a baritone singing voice (Cooksey Stage 4). This was due to participants dropping out of the choir or not achieving baritone within the 33-month study. Therefore, analyses pertaining to this data were primarily descriptive.

RESULTS

Twenty-eight participants were evaluated when singing soprano and categorized as either Cooksey Stage 0 − Prepubertal Unchanged (*n* = 15; 54%) or Cooksey Stage 1 − Mid-

Voice I ($n = 13$; 46%). The age range of all 28 participants was 8−13 years (mean = 10.2 ± 1.2) and they had a mean duration of 1.7 ± 1.1 years singing in the Boychoir (range = 0 −5 years). The participants in Cooksey Stage 0 had 0 −3 years in the Boychoir (mean = 1.37 ± 0.69); while those in Cooksey Stage 1 had 1−5 years (mean = 2.0 ± 1.41). Only 1 of 28 participants indicated enrollment in private voice lessons, which had recently begun at the time of data collection.

Of the 28 sopranos who initially enrolled in the study, only 12 returned for re-evaluation with advanced stages of vocal change due to participants who dropped out of the choir, were unwilling to return for re-evaluation, or did not physically mature to an advanced stage in the 33-month time-line of the study. Seven of the 12 re-evaluations were participants in Cooksey Stage 4 − New Voice-Baritone. To provide another comparison group for pitch breaks and perceptual measures described below, the seven participants who were Cooksey Stage 4 upon return, plus one participant who was at this stage during his initial visit, were combined into a group. The age range of this Cooksey Stage 4 group at the time of their evaluation was 13−14 years (mean = 13.4 ± 0.52), with 2−5.5 years in the Boychoir (mean = 4.06 ± 1.27).

Interrater agreement

Gwet AC1 agreement coefficients for pitch/vocal slide interval task, utilizing Altman's criterion, revealed moderate to good agreement when examining the 3 raters by pairs (Table 1). Rater consistency was considerably lower for the SSB task. Two of the raters exhibited moderate agreement, but only fair agreement existed between those 2 raters and the third rater (Table 2).

The numeric ratings for the perceptual measures of the SSB exhibited considerably higher intraclass correlations (ICC) among the 3 raters: Breathiness—ICC = 0.7227; Timbre—ICC = 0.6357; Voice Quality—ICC = 0.8723; and, Overall Vocal Quality—ICC = 0.6453.

Perceptual ratings

Aggregate descriptive statistics. The percentage of participants in which a pitch break was detected by the raters in the various pitch/vocal slide interval combinations are reported in Table 3 and are described by Cooksey Stage/ pitch/vocal slide interval (1-3-1, 1-5-1, 1-8-1). Table 4 contains the percentage of the occurrence of any pitch break in the SSB by Cooksey Stage, while Table 5 presents the rater-aggregated means for the perceptual measures of the SSB.

Pitch break potential for sopranos (Cooksey Stage 0 or 1**).** Mixed effects logistic models were fit to compare the probability of pitch break between different vocal slide intervals (Table 6).

At Cooksey Stage 0, there is always a significantly higher likelihood of pitch breaks when singing an octave interval (1-8-1) as compared to 1-3-1, and this effect becomes stronger as the starting pitch rises. Per the post-hoc analyses of specific vocal slide intervals, the likelihood of pitch breaks becomes significantly different at 1-5-1 intervals as compared with 1-3-1, but only at higher starting pitches (F4 or higher).

At Cooksey Stage 1, the likelihood of pitch breaks at 1-8-1 is significantly higher than at 1-3-1 regardless of starting pitch, but the effect is not as strong as at Cooksey 0. (Dunnett tests are unreliable here at pitches below A4 due to

TABLE 1.
Gwet AC1 Agreement Coefficients (Sample Size) for Pitch/Interval Task: Rater Pairs

	Rater 2	Rater 3
Rater 1	0.5589 (441)	0.6871 (444)
Rater 2	-	0.6011 (441)

TABLE 2.
Gwet AC1 Agreement Coefficients (Sample Size) for "Star-Spangled Banner" Task: Rater Pairs

	Rater 2	Rater 3
Rater 1	0.2783 (36)	0.2783 (36)
Rater 2	-	0.5676 (36)

TABLE 3.
Rater-Aggregated Percent Pitch Break Assessments by Cooksey Stage/Pitch/Interval Combination

Cooksey Stage	Pitch	Interval		
		1-3-1	1-5-1	1-8-1
0	G3	4.4%	0%	31.8%
	C4	4.4%	13.3%	59.0%
	F4	13.3%	33.3%	61.4%
	A4	17.8%	33.3%	66.7%
1	G3	0%	8.3%	38.9%
	C4	0%	13.9%	47.2%
	F4	0%	8.3%	47.2%
	A4	8.3%	22.2%	55.6%
4	G3	8.3%	50.0%	95.8%
	C4	50.0%	75.0%	83.3%
	F4	12.5%	33.3%	75.0%
	A4	20.1%	25.0%	83.3%

TABLE 4.
Rater-Aggregated Percent Pitch Break Assessments in the "Star-Spangled Banner"

Cooksey Stage	*Percentage*
0	51.1%
1	30.8%
4	91.7%

TABLE 5.
Rater-Aggregated Means (SDs) for Perceptual Measures (Range = 0−14.5)

Cooksey Stage	*n*	Breathiness (None-Severe)	Timbre (Dark-Bright)	Voice Quality (Simple/ Light-Hooty/Round)	Overall Vocal Quality (Poor-Exceptional)
0	15	5.60 (3.63)	7.56 (2.14)	5.85 (2.65)	6.66 (2.49)
1	13	7.66 (3.47)	7.95 (1.66)	5.31 (2.73)	6.48 (2.90)
4	8	9.83 (2.43)	8.79 (2.04)	6.50 (2.60)	5.06 (3.15)

TABLE 6.
Analysis of Interval Effects on Pitch Break Potential for Sopranos

Cooksey Stage	Pitch	Interval 1-3-1	Interval 1-5-1	Interval 1-8-1	Wald X^2	Dunnett *P* 1-3-1 vs 1-5-1	Dunnett *P* 1-3-1 vs 1-8-1
0	G3	4.4%	0%	31.8%	8.69*	0.9953	0.0064
	C4	4.4%	13.3%	59.0%	25.77‡	0.1484	<0.0001
	F4	13.3%	33.3%	61.4%	22.84‡	0.0137	<0.0001
	A4	17.8%	33.3%	66.7%	24.56‡	0.0431	<0.0001
1	G3	0%	8.3%	38.9%	9.81†	n/a	n/a
	C4	0%	13.9%	47.2%	10.26†	n/a	n/a
	F4	0%	8.3%	47.2%	12.53†	n/a	n/a
	A4	8.3%	22.2%	55.6%	19.31‡	0.1156	<0.0001

* $0.01 < P < 0.05$.
† $0.001 < P < 0.01$.
‡ $P < 0.001$.

complete absence of observed pitch breaks when singing 1-3-1 in this sample).

SSB assessments. A one-way ANOVA revealed there was no significant difference in the likelihood of pitch breaks between the two Cooksey Stages (0, 1) ($F(1,26) = 3.42$, $P = 0.0759$). Mixed effect models revealed that breathiness was significantly higher at Cooksey Stage 1 ($F(1,54) = 4.79$, $P = 0.0330$). There was no significant difference in mean Timbre ($F(1,54) = 0.45$, $P = 0.5061$), mean Voice Quality ($F(1,54) = 0.54$, $P = 0.4645$), or mean Overall Vocal Quality ($F(1,54) = 0.04$, $P = 0.8492$) between the two Cooksey Stages (0, 1) (Tables 4 and 5).

Pitch break potential longitudinal comparison (Cooksey Stages 0/1 to 4, participants with two visits). Of the eight Cooksey Stage 4 participants, seven presented with the changed voice at a return visit. (The eighth participant presented as a Cooksey Stage 4 at his initial visit and was not re-evaluated and not included in the longitudinal analysis). A given participant's initial and follow-up pitch break assessments were paired by starting pitch (G3), vocal slide interval, and individual rater. Data for starting pitch D3 used with baritones, not sopranos, was not included since it was not observed on the initial visit. Each pre-/postassessment from Cooksey Stage 0/1 to Cooksey Stage 4 was classified according to change in pitch break status which included: A] no break → no break (no change); B] no break → break; C] break → no break; D] break → break (no change). The 7 participants with two visits (first visit at Cooksey Stage 0/1 and a follow-up at Cooksey Stage 4) were also used to analyze changes in their pitch break patterns and perceptual measures on the SSB. See Table 7 for the percentage pitch break change status across all starting pitches and SSB.

Overall, pitch breaks were observed during the vocal slides (1-3-1, 1-5-1, 1-8-1) in over 53% (B + D) of assessments of participants after they progressed to Cooksey Stage 4. Pitch breaks in the SSB were observed in over 90% (B + D) of assessments following voice change to Cooksey Stage 4.

Changes in perceptual measures were evaluated using *paired t* confidence intervals on rater average scores (Table 8).

The only significant change in SSB perceptual measures was in breathiness, which increased significantly following change to Cooksey Stage 4 (mean increase = 4.60). No other measure exhibited a significant change.

DISCUSSION

This study examined the perceptual quality of voice in boys at early stages of puberty on interval vocal slides and the SSB. Interesting differences were observed in pitch breaks and some perceptual qualities between the first 2 stages identified by Cooksey (Unchanged and Mid-Voice I). This study also had a longitudinal component, which examined the change in perceptual qualities across singing tasks and

TABLE 7.
Aggregate Percentage Breakdowns of Pitch Break Change Status Across Raters for Vocal Slides (1-3-1, 1-5-1, 1-8-1) and "Star-Spangled Banner" (SSB) Assessment

Pitch Break Change Status Cooksey Stage 0/1 → 4	Vocal Slide (Overall Percentage)	SSB (Percentage)
A] no break → no break	37.5%	9.6%
B] no break → break	36.1%	57.1%
C] break → no break	9.3%	0%
D] break → break	17.1%	33.3%

TABLE 8.
Rater-Aggregated Change Scores on Perceptual Measures

Measure	*n*	Mean Change (SD)	*t*	95%CI
Breathiness	7	4.60 (3.37)	3.62*	(1.49, 7.72)
Timbre	7	0.77 (1.83)	1.11	(−0.93, 2.46)
Voice quality	7	0.32 (1.78)	0.47	(−1.33, 1.97)
Overall vocal quality	7	−1.05 (2.31)	−1.21	(−3.18, 1.08)

* $0.01 < P < 0.05$.

pubertal voice stages in boys who reached Cooksey Stage 4 (Baritone).

Pitch breaks can occur as a result of vocal instability due to untrained vocal musculature or early physiologic/maturational changes affecting vocal control. There was some variability between the soprano singers classified as Cooksey Stage 0 and Cooksey Stage 1. Specifically, there were observable differences between participants in Cooksey Stage 0 and 1 regarding pitch breaks during vocal slides. While both groups increased their number of pitch breaks as they progressed from intervals of 1-3-1 to 1-8-1 for each of the four pitches (G3, C4, F4, A4), the Cooksey Stage 1 participants began with no pitch breaks for the first 3 of 4 pitches during the 1-3-1 interval; whereas, Cooksey Stage 0 had pitch breaks for all pitches during the 1-3-1 interval. In general, participants in Cooksey Stage 0 had a greater number of pitch breaks in most intervals of vocal slides compared to Cooksey Stage 1, despite experiencing no maturational voice change. These same occurrences of pitch break occurred with the SSB. This greater number of pitch breaks for participants in Cooksey Stage 0 may be due to fewer years in the Boychoir with less experience/training in vocal performance and presumably less stability of the voice. And, while boys in Cooksey Stage 1 are starting to undergo voice change, it is assumed that they are more trained and exhibit greater control of the voice. Regardless of Cooksey Stage, the soprano singers had more pitch breaks with the octave vocal slides (1-8-1), particularly as the initial pitch rose, compared to the smaller interval vocal slides. Their limited training in vocal control hinders their ability to smoothly glide into a falsetto voice.

As would be expected, the baritone singers had higher number of pitch breaks compared to the sopranos, particularly in the octave vocal slides (1-8-1), due to their lack of development of the falsetto. We anticipated more pitch breaks in Cooksey Stage 4, compared to Cooksey Stage 0 or 1, for both vocal slide intervals and SSB, as this is a later stage of puberty when voice change is significant, and the singer is learning to sing with anatomic laryngeal changes that specifically relate to pubertal growth.

It is not surprising that perceptual qualities were similar between Cooksey Stage 0 and 1, except for breathiness which was greater in Cooksey Stage 1. The trend of greater breathiness continued as the singers experienced greater voice change and advanced in Cooksey Stages. The ability to maintain engagement/disengagement of the vocal folds during the rapid variability of pitch change in the SSB became more challenging for the baritone singers. Consequently, breathiness increased significantly, and overall vocal quality became poorer following the change to Cooksey Stage 4.

Pitch breaks are a characteristic perceptual change that indicate a young man may be transitioning through puberty. For choral directors and singing teachers, the vocal instability and lack of consistency in male voices during puberty has resulted in teachers/directors having varying opinions on the safety and efficacy of allowing boys to sing during active vocal puberty. In an effort to characterize vocal stages of change, John Cooksey developed a systematic categorization of male voices based on timbre, range, and pitch breaks. These measures have provided choral directors and voice teachers a perceptual means to begin to understand the male voice. The present study provides additional insights into specific acoustic parameters that may also be considered as the male voice changes during puberty. Specifically, pitch breaks during Cooksey's Stage 4 are noted to significantly increase during the SSB, occurring approximately 90% of the time in the singers in this study. That said, within this study, when the vocal slide interval size was smaller, there was less likelihood of pitch break, even in Cooksey Stage 4 voices. It should be noted for the purposes of this research, the SSB was performed in the same key during all testing and voice stages. Consequently, as the voice range changed, adolescent singers pitch stability decreased when attempting to sing in the original key, providing insight into the need for a change in their choral vocal role. Many of the boys were well aware at Stage 4 that they could not navigate the SSB's octave and a half interval without a pitch break. Following their initial performance, some boys asked if they could sing it again in a different key because they did not like how their voice had sounded in the original key.

The other acoustic and perceptual parameter that emerged consistently as a point of voice change and development was breathiness. Breathiness significantly increased as the boys progressed from Stage 1 through Stage 4. Breathiness was shown to be more significant than timbre and voice quality. This is in contrast to Cooksey's work which does not consider breathiness as the prominent salient feature. However, this study indicates that breathiness may indeed be a point of consideration during typical male vocal development and serve as a perceptual feature for consideration when classifying male voices. Perceived breathiness can arise from discoordination of adequate phonation and respiration and/or it may occur from glottic gapping which allows increased breath flow through the glottis during this period of physical growth. As a part of the current study, laryngeal imaging was conducted on each singer during each visit. It is beyond the scope of this paper to include all laryngeal findings. However, there is indication that the breathiness may arise from the laryngeal level (glottic closure patterns) and is not the sole result of poor vocal training or discoordination of respiration, breath support, and phonation patterns. In short, while breathiness can be lessened or perhaps eliminated through vocal training, breathiness many continue due to the physiologic nature of the larynx as a male adolescent progresses through puberty.

CONCLUSIONS

Breathiness, pitch breaks, and vocal instability can be recognized by both the male singer and those who train and conduct them. Choral directors and singing teachers should be aware that breathiness and pitch breaks may be the primary perceptual vocal cues of a young man moving through "vocal puberty" into full voice maturation. As the singers in this study were acutely aware of their perceptual changes, it may benefit the choral director/singing teacher to allow vocal flexibility and movement into a different vocal section within the ensemble to accommodate the active voice change. These accommodations (of having a singer switch parts) are difficult when maintaining a performing ensemble. However, periodic reassignments may be necessary to maintain retention of these young men so that they may to continue to sing with confidence during this tumultuous vocal period. Additionally, because each piece of repertoire will call for a different musical range, directors may need to adjust notes that reach into the extremities of the range, and should work with each student to ensure that he can comfortably reach the notes that are called for.

APPENDIX A: COOKSEY'S STAGES OF MALE CHANGING VOICE

Unchanged (Stage 0)

Range: 220−698 Hz [a3−f5] (tessitura: 277.18−493.88 Hz [c#4−a#4])

Average SF0 range: 220−260 Hz [a3−c4]

Mid-Voice I (Stage 1)

Range: 208−523 Hz [g#3−c5] (tessitura: 247−392 Hz [b3−g4])

Average SF0 range: 220−247 Hz [a3−b3]

Mid-Voice II (Stage 2)

Range: 175−392 or 440 Hz [f3−g4 or a4] (tessitura: 208−349 Hz [g#3−f4])

Average SF0 range: 196−233 Hz [g3−a3]

Mid-Voice IIA (Stage 3)

Range: 147−370 Hz [d3−f#4] (tessitura: 185−262 Hz [f#3−c4], sometimes with extension to 330 Hz [e4])

Average SF0 range: 175−185 Hz [f3−f#3]

New Voice − Baritone (Stage 4)

Range: 123−311 Hz [b2−d#4] (tessitura: 155.56−246.96 Hz [d#3−b3])

Average SF0 range: 131−165 Hz [c3−e3]

Emerging Adult Voice (Stage 5)

Range: 98−293.66 Hz [g2−d4] (tessitura: 123−416 Hz [b2−d4])

Average SF0 range: 110−139 Hz [a2−c#3]

REFERENCES

1. Cooksey J. The development of a contemporary, eclectic theory for the training and cultivation of the junior high school male changing voice. *Choral J*. 1977;18:5–14.
2. Cooksey J. *Working With the Adolescent Voice*. St. Louis: Concordia Publishing House; 1992:55–64.
3. Cooksey J. Voice transformation in male adolescents. In: Thurman L, Welch G, eds. *Bodymind and Voice—Foundations of Voice Education*. Iowa City: The Voice Care Network; 2000:718–738. 821-841.
4. Cooksey J, Beckett R, Wiseman R. A longitudinal investigation of selected vocal, physiologic and acoustic factors associated with voice maturation in the junior high school male adolescent. In: *Proceedings of a Research Symposium on the Male Adolescent Voice*. 1, Buffalo, NY: Buffalo Press; 1984.
5. Cooksey J, Beckett R, Wiseman R. A longitudinal investigation of selected vocal, physiological, and acoustical factors associated with voice maturation in the junior high school male adolescent. In: Thurman L, Welch G, eds. *Bodymind and Voice—Foundations of Voice Education*. Iowa City: The Voice Care Network; 2000:731–733.
6. Meckea A, Sundberg J. Gender differences in children's singing voices: acoustic analyses and results of a listening test. *J Acoust Soc Am*. 2010;127:3223–3231.
7. Thurman L. Boys' changing voices: what do we know now? *Choral J*. 2012;52:8–21.
8. Daw SF. Age of boys' puberty in Leipzig, 1727-49, as indicated by voice breaking in J.S. Bach's choir members. *Hum Biol*. 1970;47:87–89.
9. Juul A, Magnusdottir S, Scheike T, et al. Age at voice break in Danish boys: effects of pre-pubertal body mass index and secular trend. *Int J Androl*. 2007;30:537–542.
10. Karpati AM, Rubin CH, Kieszak SM, et al. Stature and pubertal stage assessment in American boys: the 1988−1994 third national health and nutrition examination survey. *J Adolesc Health*. 2002;30:205–212.
11. Lee P. Normal ages of pubertal events among American males and females. *J Adolesc Health Care*. 1980:26–29.

12. Tanner JM. *Growth at Adolescence*. Oxford: Blackwell Scientific Publications; 1962:156–206.
13. Tanner JM. Standards of normal growth. In: Tanner J, ed. *Foetus into Man: Physical Growth From Conception to Maturity*. London: Open Books; 1978:196–201.
14. Tanner JM. Growth and endocrinology of the adolescent. In: Gardner L, ed. *Endocrine and Genetic Diseases of Childhood and Adolescence*. Philadelphia: W.B. Saunders Co.; 1975:14–64.
15. Harries M, Walker J, Williams D, et al. Changes in the male voice at puberty. *Arch Dis Child*. 1997;77. 445−44.
16. Carskadon M, Acebo C. A self-administered rating scale for pubertal development. *J Adolesc Health*. 1993;14:190–195.
17. Norris S, Richter L. Are there short cuts to pubertal assessments? Self-reported and assessed group differences in pubertal development in African adolescents. *J Adolesc Health*. 2008;42:259–265.
18. Petersen A, Crockett L, Richards M, et al. A self-report measure of pubertal status: reliability, validity, and initial norms. *J Youth Adolesc*. 1988;17:117–133.
19. Fuchs M, Froehlich M, Hentschel B, et al. Predicting mutational change in the speaking voice of boys. *J Health*. 2007;21:169–178.
20. Hasek C, Singh S, Murry T. Acoustic attributes of preadolescent voices. *J Acoust Soc Am*. 1980;65:1262–1265.
21. Morris R, Ternstrom S, LoVetri J, et al. Long-term average spectra from a youth choir singing in three vocal registers and two dynamic levels. *J Voice*. 2012;26:30–36.
22. Pabon P, McAllister A, Sederholm E, et al. Dynamic and voice quality information in the computer phonetograms of children's voices. In: White P, ed. *Child Voice*. Stockholm: KTH Voice Research Centre; 2000:85–100.
23. Pedersen M, Moller A, Krabbe S, et al. Fundamental voice frequency measured by electroglottography during continuous speech. A new exact secondary sex characteristic in boys in puberty. *Int J Pediatr Otorhinolaryngol*. 1986;11:21–27.
24. Salomão G, Sundberg J. Relation between perceived voice register and flow glottogram parameters in males. *Acoust Soc Am*. 2008;124:546–551.
25. Schneider B, Zumtobel M, Prettenhofer W, et al. Normative voice range profiles in vocally trained and untrained children aged between 7 and 10 years. *J Voice*. 2010;24:153–160.
26. Sergeant D, Welch G. Age-related changes in long-term average spectra of children's voices. *J Voice*. 2008;22:658–670.
27. Hacki T, Heitmuller S. Development of the child's voice: premutation, mutation. *Int J Pediatr Otorhinolaryngol*. 1999;46:S141–S144.
28. Juul A, Teilmann G, Scheike T, et al. Pubertal development in Danish children: comparison of recent European and US data. *Int J Androl*. 2006;29:247–255.
29. Gwet L. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol*. 2008;61:29–48.
30. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–428.
31. Wilson JR, Lorenz KA. *Modeling Binary Correlated Responses Using SAS, SPSS and R*. 9. Springer; 2015.