Pitch, pedagogy and performance: demographic structure and vocal blending in an English cathedral boys' choir

Martin Ashley, PhD
Emeritus Professor, Edge Hill University, UK
Association of British Choral Directors

Abstract

This paper presents the results of a detailed study of the pitching accuracy and sound power level of the boys in a cathedral choir. Although cathedral choirs are expected to perform to a professional standard, they are also expected to offer a high standard of musical education to the children who sing the highest vocal part. Boys are known to take longer than girls to develop pitching accuracy yet have been the traditional choice for the top line. The paper takes as its starting point a theoretical demographic model which posits five age groups corresponding to school years within a typical choir. The pitching accuracy and sound power level of voices in each age group are examined. No significant relationship between sound power level and age was found. Boys were found to pitch accurately from the age of eleven upwards. The sometimes less accurate pitching of younger boys did not have a detrimental effect on the choir because random sharp and flat deviations tended to cancel out in the chorus effect. Issues are raised concerning the oldest boys who sing in falsetto with changed voices. These boys pitch accurately and are of value to the choir on account of their experience and repertoire knowledge. However, some singing teachers counsel against the use of the falsetto treble voice. Advances in the timing of puberty may require compensatory action to reduce the age at which accurate pitching is achieved.

Keywords boys, choristers, pitch, accuracy, blending, voice change, age, maturity

Introduction

What is a cathedral choir for? One answer that might be given is to perform sacred choral music to a high standard. Another might be the "opus dei" (literally the "work of God", i.e. the daily singing of divine office). Those who object to divine worship being any form of "concert" might see conflict here, but the question is asked for a rather different purpose. Performance, whether for God or man, is not the whole story. As the present author has argued at length elsewhere (Ashley, 2014: 18-20), cathedral choirs perform the twin functions of performance and pedagogy. Unlike a professional adult choir that might simply audition the best available sopranos, a cathedral boys' choir must bear the pedagogical burden of teaching its boys to sing, cultivating their musicianship and developing rapidly their repertoire knowledge. At one end of its age demographic it must take small boys who may only just have discovered that they have an upper register (or "singing voice"). At the other end, it must decide what to do with larger boys whose vocal functioning may be increasingly disrupted by puberty.

What might be the impact of the outer extremities of an age demographic that differs little in principle from the age demographic of a school? An ideal choir demographic structure is illustrated in Table 1 below, taken from an earlier publication by the present author (Ashley, 2014). A key argument in that publication was that for some choirs there is a threat to this ideal posed by "demographic creep". Unless there is constant recruitment of younger boys, a choir will gradually age until there comes a point where numbers are suddenly, perhaps catastrophically, decimated. Several of the oldest boys may suffer "voice break", sometimes in rapid succession[1]. It is not uncommon under such conditions to find choirs clinging to boys late into pubertal voice change, a practice that has attracted significant criticism (Trundle, 2005; Trezevant, 2007; Phillips, 2013; Williams, 2011; 2020). The issues at the younger end of the choir are of a different nature but also significant. Welch et al (1997) reported that "out-of-tune" boys outnumbered similar girls by a factor of up to 3:1. A later, large-scale evaluation of the Sing-Up programme for primary schools by the same author and colleagues (Welch et al 2009) found that girls had a statistically higher mean normalised singing assessment rating (χ = 78.97, than boys (χ = 68.82). So, there may be issues of pitching accuracy to confront with the younger boys.

Table 1
Demographic Structure of a Boy Treble Line

Year (age)	*Number*	*Function*
Year 4 (8–9)	0–2	The 'nursery'. May not yet join the choir, but the more advanced ones might become probationers to replace losses.
Year 5 (9–10)	4	On a rapid learning curve. Finding their feet and becoming the next year's mainstay of the treble section.
Year 6 (10–11)	4	Key year in which the pinnacle of development may be approached. Vital part of the young sound of the choir.
Year 7 (11–12)	4	The second key year in which musical experience is added to voices at their peak. The 'golden year'. Some voices may become leading soloists, though a minority of others may begin to pass their best.
Year 8 (12–13)	3	At least one Year 8 pupil is likely to be lost to voice change. Others may assume useful leadership roles though their range and quality of their treble voices may diminish. Boys with later puberty may still peak as soloists or leading members during this year.
Year 9 (13–14)	0–2	One or two boys for whom puberty comes later may survive as useful members of the choir, possibly with very powerful top registers—a 'super nova of the dying treble voice'.

(Ashley, 2014: 117)

[1]So-called "voice breaking" results from human agency. It is not purely a biological event as sometimes thought.

Previous research by the present author has confirmed that these problems are real. A study of seven cathedral choirs (Ashley, 2013) found that 42% of the Y8 boys (age 12:06 – 13:06) had baritone speaking voices, singing "treble" only in falsetto (and in one case not singing at all but miming). The same study also found evidence of difficulties in the youngest boys. Even in prestigious choirs able to be quite selective at voice trial, the youngest boys could not complete a test that was managed with ease by the boys in the centre demographic. Boys were asked to sing the words "You owe me five pounds" to the tune of *Happy Birthday*, with no starting note given. The primary aim of this test was to identify where the boy naturally placed his tessitura. Typical results can be heard in sound samples 1 – 3. In sample 1, a Y5 boy misjudges his tessitura with the result that the octave leap is badly mistuned and compressed. In sample 2, a Y7 boy completes the test successfully, as was almost invariably the case with the centre demographic. In sample 3, a Y8 boy clearly struggles with the conflict between his modal and falsetto voices. Similar difficulties were found with other boys whose voices were changing. They could sing known repertoire but were disturbed by a test that revealed the existence of a new voice they did not normally use.

Similar patterns have been reported by other researchers. Howard and Angus (2009) confirmed that girls pitch accurately earlier than boys. Williams et al (2009) noted an additional related difficulty. Having matched a single pitch, when asked to sustain this pitch on either the note D4 or D5 beginning quietly and increasing in loudness the youngest boys were unable adequately to complete the task. In all cases, the ability to do this was acquired during the first year of training (Williams et al, 2009). Although she found mid-puberty choristers singing successfully with a "hybrid phonation" (Williams, 2010), the same author reports on how the additional burden of performance anxiety can cause a voice that is reliably "treble" in rehearsal to crack into the baritone register during performance[2] (Williams, 2020).

Considerations such as these have led the present author to seek more information on the relative contributions made by the different stages of physical and vocal maturity that are found across a choir's age demographic. The following research questions are now asked of a cathedral boys' choir.

- How does pitching accuracy vary with age and physical maturity and how do different skill levels affect the choral ensemble?
- Is the contribution of every boy necessary for performance?
- What is the relationship between power output and number of boys, and does this vary according to the size and maturity of the boys?

Acoustic considerations

For many choirs, particularly amateur parish choirs, the replacement of older boys with "broken" voices by younger recruits presents a constant challenge. Even the more prestigious professional choirs are finding that with the perceived need to maintain both a boys' and girls' top line, the keeping of eighteen or so boy choristers threatens to

[2] As in the infamous case of David Hemmings during a performance of *Turn of the Screw* conducted by Benjamin Britten. More recently (2018), a similar misfortune befell the treble soloist during a BBC Radio 3 Choral Evensong broadcast of the Allegri *Miserere* from an Oxbridge college.

become a financial burden that some cathedral chapters might be tempted to identify as an expensive luxury[3]. There are, however, also acoustic reasons for asking the question. Bowers (1986) found the surprising result that during the sixteenth century, the age of high choral polyphony (when it might be imagined that all parts should be heard equally), there were generally significantly fewer boys relative to men than was the case during the mid-twentieth century when boys were plentiful. It may not necessarily have been the case that sixteenth century boys either over-sang or were not clearly heard, because the loudness of a sound is proportional to the logarithm of the actual power of the air compressions. Consequently, twice the number of boys does not mean twice the power of sound or, conversely, half the number of boys does not mean half the power. Ternström (1989) calculated that doubling the number of singers increases the sound level by only 3dB whereas psycho-acousticians report a difference of between 6 and 10dB to be necessary for a sound to be perceived as twice as loud (Sengpiel, 2020).

What may be more important is not the absolute number of boys but the size and power of the boys that exist and the ways they combine in ensemble - the so-called "chorus effect". Baldy notes the "enormous amount of noise" that can be produced as "rather a nice final flourish" for the final year chorister singing in a "supercharged falsetto" (Baldy, 2010: 78). Thus, boys at the top end of the age demographic might be valuable to a choir, not just for their repertoire knowledge and musicianship, but for the sheer power of their voices. Questions arise about how voices of potentially different power combine within the chorus and the extent to which individual voices, particularly those of older boys singing in falsetto, might stand out.

The combination of individual voices into a choral ensemble is an exceedingly complex matter that even now is not entirely understood. A key issue, which seems surprisingly little documented, is the nature of the single pitch that is perceived when what is being heard is any number of closely adjacent pitches. Quite how much the pitches of individual singers within an ensemble might differ from each other and deviate from the exact note pitch is clearly a fundamental issue if there are grounds to believe that the pitching accuracy of some of the singers may be outside tolerable limits. The established method for quantifying this phenomenon is to express pitch scatter (how far apart the singers are from each other) in standard deviations (SDs). Ternström and Sundberg (1988) made multitrack recordings of choir basses in rehearsal, finding pitch scatter to be 10-15 cents in trials judged by their conductor to be musically acceptable. In a later experiment employing synthesised stereophonic voices, Ternström found a result consistent with this previous work. Most listeners would tolerate a standard deviation in F0 of 14 cents, meaning that two-thirds of the ensemble would be within ±14 cents of each other. Less deviation however was preferred (Ternström , 1993). An SD of 14 cents might serve as a reference point for the present study although it is likely that singers in the soprano range might need to be closer to each other in pitch than singers in the bass range since the mistuning beats are faster and more intense at the higher pitches (Daugherty, 2005).

Pubertal Voice Change

Pubertal voice change is most commonly thought of in terms of rapid laryngeal growth identified by voice deepening, but the vocal tract grows also, not always in direct

[3] Rochester Cathedral has been the latest to announce the merging of its two former choirs (November 2019).

synchronism with the larynx. This can lead to significant disruption of higher partials (Cooksey, 2000.) with deleterious effects on timbre. The age demographic of many cathedral boys' choir can span the entire period of male puberty (Ashley and Mecke, 2013; Ashley, 2014). It is not uncommon to find a choir containing voices at each of the five stages of puberty established by J.M Tanner (Tanner and Whitehouse, 1976). This is in spite of the disapproval of some singing teachers and voice coaches who caution against singing "too high for too long" (Williams, 2011). There is clearly a loose association between physical and musical maturity. It might be expected that boys nearing the end of puberty will have greater musical maturity, but against this is the fact that rapid and asymmetric growth of the vocal folds may cause new problems of vocal control in some parts of the range.

Howard et al (1997) assessed progression in pitch matching accuracy and reported a "general singing competency for the majority" with "few children singing out of tune at the age of 11". Unfortunately, this study was only of progression in children between ages 7 to 11 and did not examine possible regression due to puberty. Yarbrough et al, (1995) proposed that if boys had not mastered pitch matching by adolescence, they might fall further behind the peer group or give up altogether as a result of the register shift they experience while going through the voice change. A preliminary investigation by these authors appeared to suggest pitch-matching ability losses between fifth and sixth and seventh and eighth grades (UK Y5-6 and 8-9). However, they could not be certain from their data that voice change was the cause.

Demorest and Clements (2007) did examine the effect of pubertal change and found little difference in perception and production scores between boys in different stages of the change. They suggested, therefore, that the register shift brought on by vocal maturation does not significantly influence pitch-matching or perception skills. However, in their study, boys were allocated to singing ranges representing comfortable tessitura for middle school boys (bass, tenor or cambiata). This might be rather different to the case of a cathedral choir where boys who might be tenor or bass in a middle school choir still sing "treble". There were other significant methodological differences. The present study compared boys on a variable scale of cents whereas a binary analysis was employed by Demorest and Clements. They considered boys "accurate" if within 50 cents of the target note. The mean cent deviation for non-accurate or missed pitches was 151.46 cents, or 1.5 semitones with a range from 51 to 530 cents. Whereas the boys in the present study were assessed note by note in an actual singing task, the boys in the Demorest and Clements study were asked to pitch match to one note sung by a male vocal model and then to pitch match only the last note of a five note vocalise sung by the model. It would seem, therefore, that there is scope for further studies similar to the one carried out by the present author.

Classification of male voice stages during puberty has been extensively researched and discussed. As a result of ongoing collaboration with a consultant paediatrician (Butler et al, 2017; Butler, 2012), the present author favours a four-phase system similar to that proposed by the German phoniatrician Michael Fuchs in his work with the Thomanerchor of Leipzig (Fuchs et al 2008). The system allows for boys for whom voice change has yet to begin through to those whose speaking voices are baritone and for whom the singing of a treble part is possible only through falsetto.

1. ***Prepubertal***. Unchanged voice, generally, below the age of eleven (Tanner 1).
2. ***Peripubertal***. Earliest signs of puberty and voice deepening, generally around ages eleven and twelve (Tanner 1 – 2).
3. ***"In-puberty***". The recognised medical threshold for full pubertal onset has been passed, characterised by a growth spurt and perceptibly deepening speaking voice, typically around ages twelve and thirteen. (Tanner 2 – 3).
4. ***"Completing-puberty***" Full maturation, the most rapid growth spurt and vocal transition to baritone quality with falsetto or "hybrid phonation" production of treble voice. Any time between the ages of thirteen and sixteen, though can occur earlier (Tanner 4 – 5).

Demorest and Clements (op. cit.) recognize phases 2 – 4 as likely to be found in a middle school choir (ages 11 – 14). Cambiata voices might be found at phases 2 and 3, whilst tenor or bass voices are possible at phases 3 and 4, although the more normal practice is to consider all such voices as "emerging baritone" in adolescent singers (Ashley, 2018).

Method

A cathedral choir in the North of England with which the author had previously worked consented to participate in an experiment designed to assess the physical maturity, pitching accuracy and power of each individual voice. Ethical approval for the study according to the Helsinki Declaration was granted by the Edge Hill University ethics committee. Parents and boys gave fully informed consent for anonymized use of the data on the understanding that the right to withdraw and review the data existed. They were offered the opportunity to hear the recordings before publication, but none took up this offer.

Vaughan Williams' *This is the Truth Sent from Above* was chosen as a test piece because it was well known to the boys who were used to singing the first verse trebles only *a capella*. All boys present first recorded the test piece in their normal singing position in the cathedral choir stalls. Then each boy recorded individually in as dry an acoustic as could be provided in the song school. For this, the boys wore headphones and recorded to a guide track consisting of a piano accompaniment with options of click track, organ, adult voice and digitally synthesised voice. In the event, the piano accompaniment proved sufficient and the other tracks were not used.

Multitrack recordings were made with Logic Pro software capturing audio at 24bit resolution and a 96KHz sample rate. The use of multitrack recording allowed different versions of the choir to be mixed at will. Although not necessary to answer the research questions about pitching accuracy and vocal power the ability to test the ensemble effects of combining different voices in different proportions was considered useful in designing further study. Shock mounted Audio-Technica AT4040 large diaphragm cardioid condenser microphones were connected to a Macbook pro computer via a Tascam US122 MkII digital interface. Response range was 20 – 2000 Hz, impedance 100 ohms, dynamic range 133dB and signal to noise ratio 82dB. When working with individual boys, a single AT4040 microphone with pop filter was used. Singer to microphone distance was maintained at 80cm, whilst microphone stand height was adjusted to match each boy's height. Gain was set at a level just below clipping for the loudest voice and held constant for all other voices.

Physical maturity profiles of each of the boys were constructed through data gathered on the day of the recording and a year earlier. The following data were gathered on both occasions:

- Age
- School year
- Height
- Weight
- Speaking voice fundamental frequency (SF0) in counting and reading tasks
- Self-identified tessitura (as described above).

Data gathered a year earlier included:

- Singing range through vocalises and pitch matching scales
- Register profile through top to bottom vocal glides
- Vocal power/dynamic range in a messa di voce exercise on the notes D4 and D5

Vocal power on the day of the recording was derived from the analysis of the musical phrases. Time did not allow the recording of glides, scales or messa di voce. Comparison of data collected on the recording day with data collected a year earlier allowed the annual growth velocity and annual voice deepening velocity for each boy to be computed and compared with the norms for the four phases of puberty.

Analysis

Each voice was analysed using the Praat software package in order to construct individual profiles of pitch deviation, power and dynamic range.

Pitch deviation

The first step in the analysis was to identify the difference in cents from sung and ideal pitches. The middle two passages of the carol (line two, *the truth of God, the God of love* G4-C5 and line three, *therefore don't turn me from your door* E4-E5) were chosen for detailed analysis of pitch deviation in individuals. The mean pitch of each settled vowel with consonant portions of the note excluded was measured for every voice and deviation in cents sharp or flat from the target pitch was computed. Target or "ideal pitch" in practice was the piano guide track. This might not be considered "ideal" in the sense that when the whole choir was singing in the cathedral each boy's reference pitch was his perception of the other boys' pitch from where he stood in the choir (see *self to other ratio* in Daugherty, 2005) . In this situation, the boys would have had the theoretical possibility of just intonation, depending on the level of attention paid to this by the choir director. Possible differences between a live *a capella* version and a version created by mixing the individual voices might be anticipated.

Boys were then ranked in order of pitching accuracy. The task was both significantly more complex and more revealing than in the case of the Demorest and Clements study as each boy's deviation from the target pitch appeared to be randomly either in the sharp or flat direction from note to note. If the mean of pitch deviation from each note were to

be used a boy who deviated both sharp and flat by a considerable amount might appear to be an accurate pitch singer since the central tendency would be close to the ideal pitch. Use of the standard deviation (SD) could result in a boy who sang mostly accurately appearing as an inaccurate pitch singer because of one or two misjudgements of an interval. For this reason, both mean and SD are shown, but ranking is in order of SD and a degree of interpretation is required.

Power and dynamic range

A sine wave tone version of the carol at 80dB reference level was created, the power of each voice being expressed relative to this. During the analysis phase, measurement was performed acoustically with an SPL meter at a fixed distance from studio monitor speakers. Gain was set to reproduce the reference tone at 80dB. The notes E5 and E4, an octave apart, represented the pitch boundaries of the vocal line *therefore don't turn me from your door* and, as might be expected, coincided with the peak and trough of the power of the boys' voices. The reference tone, by contrast, had the same power on both E4 and E5, showing 1dB less for the whole phrase because the gaps between the notes are included.

SPL for E4, E5 and the mean of the whole line was measured by this method for successively increasing numerical strengths of singers, one boy, two boys, four boys, eight boys thus:

- Deputy head chorister (the oldest boy, at the completing puberty phase)
- Head and deputy head chorister
- The four younger boys
- The four older boys
- All eight boys
- All eight boys less the one who scored less than 95% notes inside the 50 cents target identified by Demorest and Clements.

Dynamic range was computed as the difference in SPL of notes E5 and E4.

Results

Age and status

There were twelve boys present for the recording in the cathedral. Unfortunately, it was not possible to hear every boy present individually in the time available on the day. However, the eight boys so assessed represented a good age spread across the choir. Their age and status are shown in Table 2. Normal growth rate for boys not in puberty is in the region of 5-6cm pa and 2.5kg pa (Rogol et al, 2000). Velocities faster than this indicate puberty. Mean speaking voice fundamental frequencies (SF_0) of less than 200Hz also indicate puberty, whilst an SF_0 in the region of 200 to 220Hz in association with a slightly increased growth rate indicates peripubertal status. An SF_0 of less than 170Hz would sound to most ears newly "broken" with "young man" quality, whilst an SF_0 in the region of 205 – 170Hz sounds deepened though still "boyish" in timbre (Ashley et al, in preparation).

Table 2
Age and Status of Boys Recorded Individually

ID	Status	Year	Age	SF0	Growth cm/kg pa	Puberty Phase
5	D.Head	Y9	14:05	140	10.1/11.1	Completing
8	Head	Y9	13:09	176	9.1/5.9	Completing
13	Snr	Y8	13:05	165	6.8/10.7	Completing
6	Snr	Y8	13:04	190	8.4/14.5	In-puberty
2	Snr	Y7	12:08	197	3.8/4.2	Peripubertal
1	Jnr	Y6	11:02	221	3.5/3.4	Prepubertal
9	Jnr	Y6	11:01	227	5.5/2.1	Prepubertal
10	Jnr	Y5	10:01	251	5.1/4.2	Prepubertal

Three of the boys might therefore be perceived as having "broken" speaking voices, three childlike voices and two deepish boy voices. Although chronological age cannot be taken as a surrogate for pubertal status, there is a strong pattern in relation to the choir demographic. Both Y9 boys have entered the completing puberty phase, one Y8 boy has also reached this phase whilst the other Y8 boy is in puberty. The Y7 boy has reached the peripubertal phase, whilst the three boys of primary age (Y6 and Y5) are all prepubertal.

Pitching accuracy

It was readily apparent that individual singers did not behave in the same way as a whole choir. Whereas *a capella* pitch drift in a whole choir might be expected to be unidirectional (the pitch is drifting up or down and is sharpened or flattened by the end of the piece relative to the beginning) no such tendencies were found in individual singers. Deviation from piano pitch was found to be randomly sharp or flat relative to each note with little tendency for any singer to be consistently sharp or flat. Table 2 does not then show *a capella* pitch drift, it shows the ability of each boy to maintain the same pitch as the piano, ranked by Standard Deviation. SD and age are negatively correlated ($p=0.1$, $r=-0.714$, Spearman). Thus, pitching accuracy as might be expected, improves with age.

Substituting the criterion of mean for SD makes no difference to the rank order, except for singer 2. The score for this boy, a senior at the peripubertal stage, appears to have been adversely affected by one note that was 105 cents sharp. This was D5, the second note of "there" occurring at the beginning of the line *therefore don't turn me from your door*. The C5 and E5 either side were both pitched accurately (9 and 15 cents sharp respectively). Recalculation of SD with this errant note omitted produces a result of 7.1 cents, making this boy in fact the most accurate singer in the choir. SD was equal to the deputy head chorister but mean better. The least accurate singer, by either method of reckoning, was singer 1 (a junior, though not the youngest). In this case, the range from 146 cents flat to 44 cents sharp is indicative of a relatively consistent tendency to sing flat, the only such tendency found.

Table 3
Pitch Matching Capability of Boys Recorded Individually

ID	Status	Year	Age	SF0	SD (Cents)	Mean	Range # - *b*
5	D.Head	Y9	14:05	140	10.6	16	#44-31*b*
13	Snr	Y8	13:05	165	11.1	17.5	#57-16*b*
8	Head	Y9	13:09	176	12.2	17.8	#34-42*b*
10	Jnr	Y5	10:01	251	14.3	18.4	#37-59*b*
6	Snr	Y8	13:04	190	18.7	22.9	#52-66*b*
2	Snr	Y7	12:08	197	18.8	16	#105-20*b*
9	Jnr	Y6	11:01	227	19.2	22.1	#82-28*b*
1	Jnr	Y6	11:02	221	34.8	45.4	#44-146*b*

The key question for interpretation concerns what might be considered an acceptable degree of pitching accuracy. Demorest and Clements considered a boy to be accurate if he were within 50 cents (i.e. a quarter of a tone) of the target note. By this criterion, four of the boys are 100% accurate, three 95% accurate and one 74% accurate. Table 4 shows the boys ranked by age, with accuracy according to the Demorest and Clements 50 cent criterion in the right-hand column.

Table 4
"Accurate singers" according to Demorest and Clements criterion

ID	Status	Year	Age	Accuracy
5	D.Head	Y9	14:05	100%
8	Head	Y9	13:09	100%
13	Snr	Y8	13:05	100%
6	Snr	Y8	13:04	100%
2	Snr	Y7	12:08	95%
1	Jnr	Y6	11:02	74%
9	Jnr	Y6	11:01	95%
10	Jnr	Y5	10:01	95%

Here, the choir divides neatly into two. All the senior boys in Y8 and above are 100% accurate. The Y7 boy, as discussed above, would have been 100% accurate but for one aberrant note. The least accurate was a junior boy (Y6), though not the youngest. There is thus no evidence that pitching accuracy declines with either the onset of puberty or the attainment of completing puberty status, a result that agrees with that of Demorest and Clements.

Some variation amongst younger boys according to aptitude might reasonably be expected and one might expect the boys in a cathedral choir, selected for their ability, to score highly in the region of 90 – 100% pitching accuracy. However, whilst the 50 cents criterion is objective and arithmetic, it is also arbitrary, and the more subjective question

arises of whether it is good enough for a cathedral choir. Four further quantities were computed and are shown in Table 5.

Table 5

Pitch Scatter in SD (mixed voices)	13.5	
Mean pitch deviation (mixed voices)	10.8	17#-32*b* [49]
Mean pitch deviation (a capella)	22.0	95#- 3*b* [98]
Mean pitch deviation (comparator choir)	13.6	53#-29*b* [82]

The pitch scatter (i.e. how far apart the singers are from each other) for the eight voices mixed together is shown in the first row and was 13.5 cents SD. This is close to the 14 cents SD found by Ternström to be the largest value listeners would tolerate in bass voices (see above). The mean pitch deviation for the eight voices mixed was lower at 10.8 cents (range 49 cents). For the choir singing a *capella* in the cathedral, the mean pitch deviation rose to 22 cents (range 98 cents). Clearly when a piano reference pitch is available individually for each boy, the effect is to produce a more perfectly tuned version even though the mean pitching accuracy of individual voices is no higher than 16 cents at best. This would appear to be a result of blending individual voices that were randomly sharp or flat from one note to the next.

A reverberant acoustic very similar to that of the cathedral was added to the eight mixed voices. This and the *a capella* recording were played soon after the recording day to the cathedral's music director, who was asked to identify which was which. Initially, he was unable to tell the recordings apart. After several play-throughs, he correctly identified the mixed individual voices version by perceiving that the consonant placing was not quite as unified in the mixed version. He could make no distinction by pitching accuracy. Finally, a CD recording of the same *a capella* passage by a prestigious Oxbridge choir was analysed by the same method. Mean pitch deviation in this case was found to be 13.6 cents (range 82 cents). Whilst no absolute reference values can be established, it seems that a mean pitch deviation of 13.6 cents is associated with a choir of undisputed high standard whilst a mean deviation of 22 cents was also acceptable to an expert listener (the conductor of a cathedral choir). The three versions can be heard here as sound samples 4 – 6.

Sample 4: Eight voices mixed

Sample 5: All 12 boys a capella

Sample 6: Oxbridge choir

Power and dynamic range

Table 6 shows the SPL achieved by each boy for the whole verse. Mean SPL is shown in column 7 and peak SPL (invariably the note E5) in column 8. Puberty status is also shown and mass in kg in order that the size and maturity of each singer can be factored in.

Table 6
Maturity, size and power of the singers

ID	Status	Year	Age	Puberty phase	Mass kg	dB Line	dB Peak
5	D.Head	Y9	14:05	Completing	48.5	67	85
8	Head	Y9	13:09	Completing	47	62	84
13	Snr	Y8	13:05	Completing	52.5	58	78
6	Snr	Y8	13:04	In-puberty	80	59	87
2	Snr	Y7	12:08	Peripubertal	42	62	85
1	Jnr	Y6	11:02	Prepubertal	41	61	86
9	Jnr	Y6	11:01	Prepubertal	35	63	87
10	Jnr	Y5	10:01	Prepubertal	42	62	86

Statistical tests (Spearman) were carried out to identify any possible relationships between age and physical size, age and power and physical size and power. As might be expected age and physical size were positively correlated (r=0.6845). No relationships, however, were found between age and power (r=0.1667) or between physical size and power (r=-0.3988). Any notion that the largest or most physically mature boys have in general the strongest voices is not supported by these data. It is true that the deputy head chorister, aged 14:05 and completing puberty, had the strongest voice averaged across the whole phrase, but the three youngest boys, each prepubertal, were all stronger on the highest note (E5). The two largest boys, both well into puberty, had the weakest voices averaged across the phrase. Doubtless there is much of possible interest here in the way voices develop, but it is not relevant to the present research questions.

The relevance is the fact that the loudest voices on the higher notes tended to be the less accurately pitched. How would the choir fare if a separate junior choir were to be formed for younger boys who achieved, for example, a pitching accuracy of less than 95% of notes within 50 cents? Alternatively, might meeting a criterion of 95% of notes pitched within 50 cents be required to move from probationer to full chorister status?

Table 7 shows the power levels and pitch deviation resulting from different possible combinations of singers.

Table 7
Relative Power Levels and Pitch Deviation for Combinations of Singers

Note	1 boy (8)	1 boy (5)	2 boys (5+8)	4 boys (oldest)	8 boys	4 boys (youngest)	7 boys
SPL Line (dB)	58	61	65	67	72	66	72
SPL peak (dB)	71	76	79	84	85	81	85
MPD (cents)	16.9	14.4	15.1	9.2	10.8	12.8	11.8
SDP (cents)	12.4	11.6	8.9	7.3	7	8	9.5

Reading along the top row, the two head choristers together produced an SPL of 65dB averaged over the whole of the line "*therefore don't turn me from your door*". Doubling this number by adding the other two most senior boys resulted in a gain of 2dB. Doubling

again to all eight boys produced a further gain of 5dB. The four youngest boys produced 1dB less than the four oldest boys or 1dB more than the two head choristers These results are broadly in line with Ternström's calculation that doubling the number of singers produces about 3dB of gain in sound pressure level, or, more precisely, that the addition of two sound sources of equal power results in a 3dB increase. If a gain of 6 and 10dB results in the perception that the sound is twice as loud all eight boys singing together should therefore be perceived as a little more than twice as loud as the deputy head chorister singing on his own.

Potentially more informative is the observation that the removal of the one boy who pitched only 74% of the notes to within 50 cents resulted in a difference in SPL of less than 1dB, both for the whole line and for the highest note only (middle row of the table). It might be concluded from this kind of observation that the boy's off pitch singing can be accommodated by the choir. However, when the pitching accuracy for the different combinations is examined (third row) it appears that adding voices that individually might be to varying degrees inaccurate improves the overall pitching accuracy of the chorus. Thus, the head chorister on his own has a mean pitch deviation of 16.9 cents but adding the other three boys in Y8 and Y9 improves this to 9.2 cents MPD. The four youngest boys achieve 12.8 cents, better than either of the head choristers, either on their own or in combination. Seven boys together achieved 11.8 cents MPD, whilst adding the least accurate singer (MPD 45.4 cents) actually improved the overall result by 1 cent. The explanation for this that most readily presents is found in the randomness of each boy's individual deviation from flat to sharp from one note to the next. The random deviations result in a converging central tendency as a curious and remarkable feature of the chorus effect that demands further research.

Discussion and conclusion

A defining feature of the English cathedral system is the way in which children (boys and girls separately or combined) sing with adults in choirs which to all intents and purposes can be considered professional. A logical and reasonable assumption might be that the children's voices, which are intensively trained, should reach a similar standard of pitching accuracy to the adult voices. There is little point in paying professional adult altos, tenors and basses if the ensemble is spoiled by children whose pitching accuracy is noticeably inferior. This study asked whether pitch matching skills vary with age, experience and physical maturity. It has shown that they do. Perhaps unsurprisingly, the youngest boys in a choir pitched less accurately than the older ones. More usefully, the study showed that the standard achieved by the oldest boys at age 13 or 14 was comparable to adults, achieved by the age of 11 or 12 and, once achieved, did not deteriorate. Moreover, it showed that boys aged 10 or 11 were close to this standard and that individual pitching inaccuracies tended to cancel out in the chorus effect. Even a boy who individually fell below an arbitrary standard of 95% of notes within 50 cents could be accommodated by the chorus effect.

Superficially, the traditional cathedral model of five age groups is validated for the way in which it results in a choir of good performing standard in which children can sing perfectly well alongside adults. It appears equally validated in the pedagogical model it offers of young boys learning to sing accurately through being placed amongst older boys

in a performing context. The study also asked whether pubertal voice change impacted upon pitching accuracy and found that it did not. Boys of 13 or 14 years of age who were clearly quite advanced in puberty sang accurately and did not have unduly loud voices that failed to blend. Indeed, the younger and older boys appeared to complement each other through having strengths in different parts of the register. Although it was not directly demonstrated by the study, the assumption that boys of 13 or 14 have a maturity level that makes them useful leaders and confident soloists within the choir also validates the traditional cathedral model. It was the case that two boys of completing puberty status held positions of responsibility that appeared uncompromised by any decline in pitching accuracy anywhere in the register at pitches required by the test piece. The model is therefore again validated. Were English boys to sing alto and descend regularly below the range of the test piece (as is the case in Germany), the result might be different.

There are, nonetheless, threats to this model at both ends of the demographic. In the case of older boys, a growing body of opinion amongst singing teachers has questioned practices such as a boy of over 14 years of age, with a speaking voice pitch of 140Hz who has gained over 10cm height and 11kg weight within a year continuing to sing "treble". In the case of younger boys, concern has been expressed both in the UK and in Germany about the poor quality of music education and singing training in primary and nursery schools and, indeed, the home (OFSTED, 2012; Eckart, 2010). Given what is known through the work of Graham Welch and his colleagues, cathedral music directors must be fishing in a pool of primary school boys where the numbers capable of even an entry standard of pitching accuracy can be hard to find. This study has indicated perhaps more objectively than in the case of a subjective voice trial possible regions where such a standard might lie. The criterion of 75% of notes within 50 cents looks promising.

The issue is given urgency by the downward shift in puberty timing that appears to be happening. The present author has written extensively about this elsewhere, maintaining a relatively sceptical stance regarding some of the more exaggerated claims that appear in the popular press and social media. Nevertheless, puberty timing in boys does seem to have advanced sufficiently to encroach significantly upon the 13 and 14-year-old age group (see Ashley and Mecke, 2013 for a detailed analysis). The 13 and 14-year-olds in the present study were all in puberty, three of them having reached the relatively advanced phase of completing puberty. If singing teachers continue to press their case that there is a possible threat to vocal health and a probable disadvantage to future adult singing technique, cathedral choirs are going to be increasingly reluctant to retain such boys, as some already are (Ashley, 2013).

To these considerations must be added the disruption of providing for girls as well. There can be no going back on this long overdue development and certainly no justification for denying an equal opportunity to young girls. However, setting aside the well-known aversion amongst many boys to joining choirs, it must also be considered that young girls' pitch-matching abilities are generally superior to those of boys. If boys are also going to need help because by the age of 13 most of them no longer have suitable voices, cathedral choirs may yet need to review their traditional model despite it surviving the interrogation of this study. Future research might well develop further the pioneering techniques of Jers and Ternström (2005) or Zadig (2011), who have used the technique of recording every singer simultaneously with a separate microphone. This kind of technique remains in its infancy (Howard, 2020) but the present study points towards its

considerable potential in exploring pitching accuracy where just intonation prevails in *a capella* singing, perhaps by boys and girls together.

References

Ashley, M., Butler, G. and Howard, D. in preparation. Longitudinal sampling of the adolescent male speaking voice as a component of developmental assessment in general medical practice.

Ashley, M. 2018. "What voices have emerged? Lessons on boys' vocal dispositions and choral tone from a new choral leaflet series." *Music Education Research* 21 (2): 135-149. DOI: 10.1080/14613808.2018.1534819

Ashley, M. 2014. *Contemporary Choral Work with Boys*. Oxford: Compton

Ashley, M. 2013. "The English choral tradition and the secular trend in boys' pubertal timing." *International Journal of Research in Choral Singing* 4 (2): 4-27.

Ashley, M. and Mecke, A-C. 2013. "'Boyes are apt to change their voices at about fourteene yeares of age': an historical background to the debate about longevity in boy treble singers." *Reviews of Research in Human Learning and Music*. 1 1-19.

Baldy, C. 2010. *The student voice: an introduction to developing the singing voice*. Edinburgh: Dunedin.

Bowers, R. 1986. "The Vocal Scoring, Choral Balance and Performing Pitch of Latin Church Polyphony in England, c. 1500-58." *Journal of the Royal Musical Association* 112 (1 [1986 – 1987]): 38-76.

Butler, G. 2012. *Tanner Stages v Puberty Phases*. University College Hospital Clinical Education Working Group, 7th. February.

Butler, G., Howard, D. and Ashley, M. 2017. A novel iPhone app: "Speech Test©" is a non-invasive way to determine the onset of puberty in boys. Poster presentation to International Society of Paediatric Endocrinologists.

Daugherty, J. 2005. *Choir Acoustics: an empirical approach to the sound you want*. Division of Music Education & Music Therapy, the University of Kansas.

Demorest, S. and Clements, A. 2007. "Factors influencing the pitch-matching of junior high boys." *Journal of Research in Music Education* 55 (3): 190-203.

Eckart, E. 2010. "Auf einen Ton mit dem Thomanerchor." *Merian Sachen*, retrieved from http://www.merian.de/magazin/leipzig-aufeinen-ton-mit-dem-thomanerchor.html

Fuchs, M., Froehlich, M., Hentschel, B., Stuermer, I., Kruse, E., Knauft, D. 2007. "Predicting mutational change in the speaking voice of boys." *Journal of Voice* 21 (2):169-78.

Howard, D. and Angus, J. 2009. "A comparison between singing pitching strategies of 8 to 11 year olds and trained adult singer." *Logopedics, Phoniatrics, Vocology* 22 (4): 169-176. https://doi.org/10.3109/14015439709075331

Howard, D., Angus, J. and Welch, G. 1994. "Singing pitching accuracy from years 3 to 6 in a primary school." *Proceedings of the Institute of Acoustics* 16 (5): 223-230.

Jers, H. and Ternström , S. 2005. "Intonation analysis of a multi-channel choir recording." *TMH-QPSR* 47 (1): 001-006.

Kahlin, D. and Ternström, S. 1999. "The Chorus Effect Revisited - experiments in frequency-domain analysis and simulation of ensemble sounds." Proceedings of Euromicro '99, Milano, Italy, September 8-10. IEEE Computer Society 1089-6503/99 (2): 75-80.

OFSTED 2012. *Music in Schools, wider still and wider: quality and inequality in music education 2008 - 2011*. Office for Standards in Education report 110158. Manchester: DfE.

Phillips, K. 2013. *Teaching Kids to Sing*. Boston MA: Schirmer. pp.90-94.

Rogol, A., Clark, P. and Roemmich, J. 2000. "Growth and pubertal development in children and adolescents: effects of diet and physical activity." *American Journal of Clinical Nutrition* 72(suppl): 521S–8S.

Tanner J., and Whitehouse, R. 1976. "Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty." *Archives of Disease in Childhood* 51:170-9.

Ternström, S. and Sundberg, J. 1989. "Formant frequencies of choir singers." *Journal of the Acoustical Society of America*, 86 (2): 517-522.

Ternström, S. 2003. "Choir acoustics: an overview of research published to date." *International Journal of Research in Choral Singing* 1 (1):3–12

Trezevant, F. 2007. Personal communication.

Trundle, D. 2005. *Changing voices: An approach to adolescent voice training*. Milton Keynes: Voicesource.

Welch, G., Sergeant, D. and White, P. 1997. "Age, sex and vocal task as factors in singing 'in tune' during the first years of schooling." *Bulletin of the Council for Research in Music Education* 127: 155-162.

Welch, G., Himonides, E., Saunders, J. Papageorgi, I., Vraka, M. Costanza, P. and Cynthia Stephens, C. 2009. *Researching the second year of the National Singing Programme in England: An ongoing impact evaluation of children's singing behaviours and singer identity*. London: iMerc.

Welch, G. 2006. Singing and Vocal Development. In: G.McPherson (ed.), *The Child as Musician.* New York: Oxford University Press. pp.311-329.

Williams, J., Welch, G. and Howard, D. 2009. "An exploratory baseline study of boy chorister vocal behaviour and development in an intensive professional context." *Logopedics Phoniatrics Vocology* 30: 158-162

Williams, J. 2010. *The implications of intensive singing training on the vocal health and development of boy choristers in an English cathedral choir.* Unpublished PhD thesis, Institute of Education, London.

Williams, J. 2011. Teaching young voices safely. Paper presented at the Foundations for Excellence, Dartington Hall.

Williams, J. 2020. "Boiling frogs and tipping points: Which sung pitch range is best for boys during voice change?" *ABCD Choral Singing Research*, 1(1):

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

Zadig, S., 2011. The single voice in the choral voice: how does the role of informal leaders in the choral voice affect the other singers? In. U. Geisler & K Johansson (eds) *Choir in Focus 2011*. Malmo: Kör Centrum Syd.