Article:

D'Amario, S., Howard, D.M., Daffern, H. et al. (1 more author) (2020) A longitudinal study of intonation in an a cappella singing quintet. Journal of Voice, 34 (1). 159.e13-159.e27. ISSN 0892-1997

https://doi.org/10.1016/j.jvoice.2018.07.015

A Longitudinal Study of Intonation in an *a cappella* Singing Quintet

*Sara D'Amario, †David M. Howard, *Helena Daffern, and ‡Nicola Pennill, *Heslington, †Egham, and ‡Sheffield, United Kingdom

Summary: Objective. The skill to control pitch accurately is an important feature of performance in singing ensembles as it boosts musical excellence. Previous studies analyzing single performance sessions provide inconclusive and contrasting results on whether singers in ensembles tend to use a tuning system which deviates from equal temperament for their intonation. The present study observes the evolution of intonation in a newly formed student singing quintet during their first term of study.

Methods/Design. A semiprofessional singing quintet was recorded using head-worn microphones and electrolaryngograph electrodes to allow fundamental frequency (f_o) evaluation of the individual voices. In addition, a camcorder was used to record verbal interactions between singers. The ensemble rehearsed a homophonic piece arranged for the study during five rehearsal sessions over four months. Singers practiced the piece for 10 minutes in each rehearsal, and performed three repetitions of the same pieces pre-rehearsal and post-rehearsal. Audio and electrolaryngograph data of the repeated performances, and video recordings of the rehearsals were analyzed. Aspects of intonation were then measured by extracting the f_o values from the electrolaryngograph and acoustic signal, and compared within rehearsals (pre and post) and between rehearsals (rehearsals 1 to 5), and across repetitions (take 1 to 3). Time-stamped transcriptions of rehearsal discussions were used to identify verbal interactions related to tuning, the tuning strategies adopted, and their location (bar or chord) within the piece.

Results/Discussion. Tuning of each singer was closer to equal temperament than just intonation, but the size of major thirds was slightly closer to just intonation, and minor thirds closer to equal temperament. These findings were consistent within and between rehearsals, and across repetitions. Tuning was highlighted as an important feature of rehearsal during the study term, and a range of strategies were adopted to solve tuning related issues. This study provides a novel holistic assessment of tuning strategies within a singing ensemble, furthering understanding of performance practices as well as revealing the complex approach needed for future research in this area. These findings are particularly important for directors and singers to tailor rehearsal strategies that address tuning in singing ensembles, showing that approaches need to be context driven rather than based on theoretical ideal.

Key words: Intonation−Pitch drift−Tuning−Singing ensemble−Ensemble communication.

INTRODUCTION

Tuning is an essential characteristic of good choral singing practice, at the forefront of critical reviews, director's manuals, and singing tutors.[1] Beyond the importance of pitch matching, whereby singers produce accurate unison singing within their respective parts, in *a cappella* part singing there is the additional issue of which tuning systems and temperaments should and are employed for a group to be "in tune." There are different ways to consider tuning in singing ensembles and pitch drift is a topic of common interest to researchers and practitioners alike (see Havrøy[2] for a discussion of the complex tuning issues for *a cappella* singing groups).

Accepted for publication July 11, 2018.

Declarations of interests: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

From the *Department of Electronic Engineering, University of York, York, United Kingdom; †Department of Electronic Engineering, Royal Holloway, University of London, Egham, Surrey, United Kingdom; and the ‡Department of Music, University of Sheffield, Sheffield, United Kingdom.

Address correspondence and reprint requests to Sara D'Amario, Department of Electronic Engineering, University of York, York Y010 5DD, United Kingdom. E-mail: Sda513@york.ac.uk

Journal of Voice, Vol. 34, No. 1, pp. 159.e13−159.e27

0892-1997

https://doi.org/10.1016/j.jvoice.2018.07.015

Empirical research in this area, whilst sparse, has focused on different perspectives of choral tuning including predictions of pitch drift, pitch drift in performance, and perception preferences for different intonation systems.[3] Investigating tuning practices in *a cappella* part singing, Devaney et al[4] found no evidence of pitch drift in an exercise written by Benedetti in the sixteenth century to illustrate potential pitch drift associated with "pure tuning," when performed by four expert 3-part ensembles. They hypothesized that this was due to the shortness of the exercise and the likelihood of retaining a pitch memory for the start of the piece throughout the eight-bar excerpt.

Exploring predicted pitch drift in three especially composed pieces, Howard[5] found that when modulation occurred even over a very short piece, in a single performance by one quartet, the singers had a tendency to drift in pitch. He also found that an exercise composed for the study, named "Exercise 3", was most suitable for measuring pitch drift as it avoided use of a seventh chord. In two performances of the same piece from the prior study, sung by a different quartet comprising music students, it was found that the singers drifted beyond the just intonation prediction and a long way far from equal temperament.[6]

Using the choral synthesis system described in Howard et al,[7] the tuning of student singers was analyzed as they replaced either the alto or soprano line of "Exercise 3" when listening to the other three parts tuned either in equal temperament or just intonation over headphones.[8] Singers produced less stable pitches across tones and were more "out of tune" when tuning with the justly tuned rather than the equally tempered version of the synthesis, implying that the singers in this study tended towards equal temperament.

Devaney and Ellis[8] highlighted the importance of considering both vertical and horizontal tuning, proposing an approach to account for both, which utilizes automated f_o extraction and machine learning combining theories of sensory consonance and tonal attraction respectively.

Analyzing interval sizes performed by four expert 3-part vocal ensembles, Devaney et al[4] found that whilst most vertical and horizontal tuning was in line with equal temperament both minor and major thirds varied in tuning with examples of intonation close to equal temperament, just intonation and Pythagorean tuning. Major 6ths were found to be consistently tuned to equal temperament.

The interval of a third has received particular scrutiny in research from a performance and perception perspective due to the large discrepancies between tuning systems (particularly equal temperament and just tuning) for these intervals, with a general understanding amongst trained *a cappella* ensembles to narrow major thirds and widen minor thirds to be more in keeping with just intonation (Potter[9] p. 160). Mayer[10] in describing how choirs can aspire to just tuning asserts that it is "certainly the most difficult of all intervals to sing in tune" (p. 110).

Focusing on the tuning of thirds but from the perspective of the listener, Ternström and Nordmark[11] conducted a study on tuning perception whereby expert listeners (mainly choral musicians but some orchestral) tuned synthesized dyads into major thirds. The mean results were closer to equal temperament than just intonation; however, the spread of results within subjects suggested no preference to a particular tuning system. Listeners distinguished between equal temperament and just intonation in another perceptual study, which used synthesized sounds to consider pitch drift in short chord progressions, however preference to tuning system was found to be individual.[12]

The results emerging from empirical research which reveal ambiguous perception preferences towards specific tuning systems/temperaments are also reflected in literature discussing best practice, in which the issue of tuning and temperament in *a cappella* singing continues to be highly topical and often contentious. In "A Performer's Guide to Renaissance Music" Planchart[13] asserts that, "Given the tenacity of the resistance of modern singers to just intonation," singers will find it difficult to deviate from equal temperament but "directors should ultimately neither give up or let up" (p. 38). In the same book, Blachly,[14] also extolling the importance of just intonation as producing "a more satisfying in-tune result," acknowledges that "training a small choir or vocal ensemble to sing in tune can be the most difficult challenge facing the director of the early music ensemble" (p. 25).

The application of just intonation as common practice has been an area of dispute for some time, although it is often purported to be the ideal practice for *a cappella* singing ensembles, especially when performing early music. Barbour[15] insists that "there is no system of tuning that has the virtues popularly ascribed to just intonation. Neither singers nor violinists use just intonation" (p. 48) whilst Timm[16] comments that "A Cappella choirs and string quartets [...] often boast of the use of just, or true, intonation instead of the tempered scale" (p. 19).

More recently, it has become commonly reported as a trait of "good ensemble singing" for professional groups to employ just tuning: "Performances by vocal groups such as The Hilliard Ensemble, The Tallis Scholars, and Gothic Voices have made it apparent that approaching perfection in tuning is not an impossible dream" (Duffin,[17] p. 287).

In addition to the theories and practice of the tuning of *a cappella* performances is the issue of intent and the extent to which groups actively work towards a specific tuning system and how they go about achieving their goal. Work on tuning has been observed to be a consistent feature of rehearsals of professional *a cappella* vocal ensembles[18]; however, there has been little research which focusses on specific ensemble rehearsal strategies for tuning and their evolution over a series of sessions.

Observational studies of small ensembles have demonstrated ways in which preparation for performance requires musical and social coordination, generally achieved through a framework of rehearsals and performance goals, with variation between groups of different type, size and familiarity.[19–22] As part of a study of ensemble rehearsal approaches, Chaffin and Imreh[23] categorized rehearsal tasks as "basic," "interpretive," "expressive," and "strategic." This framework was later adapted and applied in studies of ensemble rehearsals, including that of Ginsborg et al,[24] a longitudinal study of rehearsal of a professional voice and piano duo. Using verbal utterances to track the focus of the rehearsals, they characterized work on pitch and intonation as "basic" musical dimensions. Over the course of the study they observed a shift from these more "basic" tasks in early rehearsals to a greater emphasis on "interpretive" tasks (such as expressive intentions) in later sessions. This framework was also used to explore differences in rehearsal approaches in a small-scale study (four duos) of newly-formed and established student and professional ensembles.[25] There were no differences found in verbal utterances referring to "basic" musical dimensions relating to expertise or familiarity, although all participants mentioned pitch.

Studies of rehearsal techniques and performance practices addressing the issue of tuning in *a cappella* singing groups are scarce. The limited studies employing empirical methods have so far been inconclusive and, when investigating performance trends, have generally been based on single performance sessions rather than repeated takes. This paper provides a novel contribution to research in this area by introducing a mixed method repeated measures design across several rehearsal sessions in a newly formed *a cappella* vocal quintet ensemble. It combines quantitative performance data with observational frameworks of the verbal interactions of the group during rehearsals to allow for analysis of tuning in relation to practice sessions, addressing the following research questions:

1. Horizontal tuning: Does the singing quintet produce a pitch drift representative of just intonation predictions or maintain horizontal tuning in equal temperament?
   a. Do these horizontal tuning trends change pre-rehearsal and post-rehearsal?
   b. Do these horizontal tuning trends change longitudinally over rehearsal sessions spanning four months?
2. Vertical Tuning: Does the singing quintet tune thirds within chords towards just intonation or equal temperament?
   a. Do these vertical tuning trends change pre-rehearsal and post-rehearsal?
   b. Do these vertical tuning trends change longitudinally over rehearsal sessions spanning four months?
3. How do group members address tuning issues in rehearsals, as observed in their verbal interactions?

METHOD

Participants

Ethical approval for the study (with reference D'Amario070817) was obtained from the Physical Sciences Ethics Committee (PSEC) at the University of York (UK). A newly formed soprano, mezzo, mezzo, tenor, and bass singing quintet was recruited for the study (3 females, age Mdn = 23, Range = 6). Singers were postgraduate students in ensemble singing at the Department of Music of the University of York. The ensemble became established as a regular quintet working towards performances and Masters exams. They had formal coached rehearsals once a week, and additional regular rehearsals throughout the duration of the study in preparation for their final exams. All musicians had extensive experience performing in choir (Mdn = 10.8, Range = 11) and formal singing training with a professional singing teacher (Mdn = 8, Range = 13). They reported that none of the singers had absolute pitch.

Materials

The chorale "Jes, mein Hort und Erretter" from the Cantata BWV 154 "Mein liebster Jesus ist verloren" composed by Johann Sebastian Bach, and arranged for the singing ensemble in the study by the first author, was used for the analysis of the evolution of tuning across rehearsal. This piece was also used in a parallel study investigating the developmental aspects of synchronization in the same singing quintet. [26]The original Bach chorale was arranged avoiding repeated notes and limiting semitones, to facilitate tuning analysis based on f_o tracking (see Section Analysis). Tuning for each note can be potentially difficult to calculate in the f_o signal when melodies move chromatically, since the expected *vibrato* range for classical singers might span a semitone and therefore it would be difficult to detect each note. Similarly, tuning in repeated notes during *legato* singing can be difficult to analyze, if singers do not produce a noticeable pause in phonation between notes. A piece with such attributes, maximizing tuning analysis, was difficult to find, and arrangement of the piece was preferred. The arranged piece features 6 *legato* phrases performed to the vowel /i/. The piece presents a clear homophonic structure with a stable rhythm, and simultaneous entries and breaths, as shown in Figure 1. Expressive markings were not given in order to investigate aspects of rehearsal, including tuning that might emerge spontaneously.

Apparatus

Singers wore head-mounted close proximity microphones (DPA 4065) placed on the cheek of the singer at approximately 2.5cm from the lips. Stereo recordings of the repeated performances were collected using a stereo condenser microphone (Rode NT4). The latter was placed at equal distance in front of the singer at approximately 1.5m from the lips. Singers also wore electrolaryngograph electrodes (Lx) from Laryngograph Ltd. (www.laryngograph.com), placed on the neck at the level of the larynx, and kept in place with an adjustable strap. Lx is a non-invasive, widely used method for the analysis of the singing voice.[27] It has been recently used to investigate several aspects of singing ensemble performances, such as synchronization,[26,28,29] blending,[30] and tuning,[5,6] as it allows the identification of the individual contribution of each singer. Each Lx was attached to a preamplifier (ART CleanBox Pro) to reduce noise and interference over long cable runs. The 12 outputs (5 Lx with preamplifiers, 5 head-mounted microphones, and the stereo microphone with right and left channel) were connected to a multichannel audio interface (Focusrite Liquid Saffire 56) connected to a PC. The 12 outputs were then recorded in synchrony using a digital audio workstation (Reaper 5.40), set at 24-bit depth and 44.1kHz sampling frequency. Rehearsals were video-recorded with a tripod-mounted video camera (Sony MV1 Music Video recorder), with a unidirectional 120 degree XY stereo microphone. The experiment took place in a recording studio of the Department of Electronic Engineering at the University of York; the room was treated with absorptive acoustic material.

Design

This investigation is a longitudinal study consisting of five rehearsal sessions based in laboratory. The above piece was practiced for approximately 10 minutes during each rehearsal. Three repeated performances of the pieces were recorded pre-rehearsal and post-rehearsal. The Lx and audio recordings of a total of 30 repeated performances of the clearly homophonic piece were collected across the five rehearsals. The entire laboratory sessions were video recorded, to minimize the attention on the camera.

An additional piece, mostly contrasting in rhythmical content compared with the previous clearly homophonic piece, was also used for the study to investigate interpersonal synchronization between musicians in relation to the complexity of the piece rehearsed. Synchronization is out of the scope of this paper, and the results are reported in

FIGURE 1. Piece from a previous study[26], showing the major and minor thirds, highlighted with arrows and brackets respectively, that were selected for the analysis of vertical tuning. The full data set of notes was used for the analysis of horizontal tuning. The figure is ©the authors, licensed CC-BY

D'Amario et al.[26] Singers were invited to rehearse the more complex piece for 10 minutes, and performed three repetitions pre-rehearsal and post-rehearsal, as with the clear homophonic piece. The order of recording and rehearsing the two pieces was randomized within rehearsals.

Procedure

The five laboratory sessions took place over a four-month period, from September 2017 to January 2018. Prior to the first session, participants filled in a background questionnaire and gave written consent form. The first four sessions were approximately 2.5 weeks apart from each other, as shown in Table 1. The fifth lab session was originally planned three weeks after the fourth session, which was two days before their Masters exam. Due to illness and Christmas break, the exam was postponed until eight weeks after the fourth rehearsal, and the fifth lab session took place two days before the public performance. During the lab session, the quintet stood in a semi-circle of approximately 3m in diameter, in

TABLE 1.
Rehearsal Sessions Across a 16-Week Period and Allocation of Time to Tuning

						Total
Rehearsal number	1	2	3	4	5	
Week	1	3	5	8	16	
Rehearsal duration (sec)	712	315	770	618	778	3193

the sequence soprano, mezzo, mezzo, tenor, and bass. Each laboratory session lasted approximately 1 hour.

Singers were not aware of the purpose of the study. In order to encourage a natural approach to rehearsal, the group were asked to create an "expressive performance" of the pieces, which had no performance directions. During the rehearsal periods the researchers left the room, and the singers were asked to work freely on the piece however they chose. Singers only rehearsed the pieces during the five lab rehearsals, and for this reason the score was retained by the first author at the end of each lab session. This was implemented so the authors could record changes in the tuning of the given piece that evolved during the first term of study. A reference pitch A3 was given on a diapason before the three repeated performances recorded pre-rehearsal and post-rehearsal. The quintet was free to set their own tempo.

Analysis

Three aspects of tuning were analyzed: i) horizontal tuning, ii) vertical tuning, and iii) rehearsal strategies used during the lab rehearsal in relation to tuning, as shown in Table 2. In order to investigate horizontal and vertical tuning, the f_o estimates in Hertz and the corresponding timestamps with a time step of 1 millisecond were extracted from the Lx and audio recordings based on the head-mounted microphone, using Praat.[31] The two sets of data of each recording were imported into Excel as a tabular list of data. An automated peak-picking algorithm, TIMEX,[28] was used to extract the note beginnings and endings of each note from the acoustic and Lx data imported in Excel, and a macro was then implemented to compute the average frequency in Hertz of each note. This algorithm, tested on a set of singing duo recordings, proved to be a valuable and successful method for onset and offset detection in ensemble singing.[28] The data extraction automated through TIMEX was then visually cross-validated by the first two authors (SD and DMH). Notes at which pitch errors occurred, due to signal processing issues (ie, weak signal) or the singers performing wrong notes (ie., featuring a measured deviation from the expected ET value greater than 130 cents), were 1.9% of the full data set. They were identified comparing Lx and audio recordings with the notated scores, and were excluded from the analysis.

In order to analyze the pitch drift during each of the performances, a reference set of f_o is required for the tuning systems of interest; in this case, equal temperament and just temperament.[5,6] These reference f_o were calculated as frequency multipliers to the tonic of the key of the chorale (see Figure 1), which is F as it is in F major, and the starting note of the tenor part (F3) was selected. The procedure for calculating the equal temperament ratios involved multiplication (division) by the twelfth root of two to move up (down) by a semitone. The procedure for calculating the just ratios has two steps: (a) within a chord the intervals are calculated using integer harmonic ratios depending on the interval (eg, a fifth is 3/2, a major third is 5/4, a minor third is 6/5, etc.), and (b) chord to chord where a search is carried out to find the nearest harmonic ratio between one of the notes of each of the chords in the following order: unison, octave, fifth, fourth, major third, minor third, etc. Further details on tuning systems and frequency ratios can be found in Howard and Angus.[32] The measured f_o values are entered into the spreadsheet, and the f_o of each sung note is divided by the measured f_o value for the first note (F3) of the tenor part which is the reference note for the analysis as indicated above. To establish how close the sung notes were to equal temperament or just temperament, the measured frequency ratios are divided by the equal (just) tempered ratios. For the analyses presented below, the results have been converted to cents (1 cent is one hundredth of a semitone) to enable comparisons to be made.

The horizontal analysis was based on the whole set of notes (ie, 42 notes) included in the piece. A total of 15 major thirds and 23 minor thirds across parts were selected for the vertical analysis, as shown in Figure 1. The thirds were simple intervals, except for one compound major third, between bass and tenor in note n 42, which was also selected for the analysis. This interval was considered relevant to the analysis of thirds, being

TABLE 2.
Aspects and Parameters of Tuning Investigated and Corresponding Recordings and Dataset

Aspect	Parameter	Recordings	Dataset
Horizontal deviation	Pitch drift, and tuning consistency and dispersion	Lx and audio	Deviation for each note/singer/repetition
Vertical deviation	Tuning stability, consistency and dispersion	Lx and audio	Deviation for major and minor thirds
Rehearsal strategies	Time spent on tuning and approaches to tuning	Video	Rehearsal episodes

TABLE 3.
Multilevel Linear Models Implemented in the Study

Response Variable	Primary Fixed Effects	Nested Fixed Effects	Random Effect	Data Set
Drift	Rehearsal number	Stage	Take, note and singer number	All notes
Consistency	Rehearsal number	Stage	Take and singer number	All notes
Dispersion	Rehearsal number	Stage	Take and singer number	All notes
Drift	Rehearsal number	Stage	Take and interval number, singer pair	Major thirds
Consistency	Rehearsal number	Stage	Take number	Major thirds
Dispersion	Rehearsal number	Stage	Take number	Major thirds
Drift	Rehearsal number	Stage	Take and interval number, singer pair	Minor thirds
Consistency	Rehearsal number	Stage	Take number	Minor thirds
Dispersion	Rehearsal number	Stage	Take number	Minor thirds

the last chord of the piece. Three metrics of horizontal and vertical tuning were measured as follows:

- Pitch drift, as indexed by the pitch deviation from ET and just intonation
- Tuning consistency, as indexed by the SD of measured deviations computed for each repetition (ie, take) pooling the 42 notes or the selected thirds to analyze horizontal or vertical consistency, respectively
- Tuning dispersion, as indexed by the range of measured deviation computed across notes or selected thirds for each repetition, similarly to the procedure implemented for tuning consistency analysis

Multilevel linear-models of the response variables (ie, f_o deviation from predicted values, SD and range of measured deviation) were then implemented to test the primary fixed effects of rehearsal, and the fixed effects of rehearsal stage (ie, pre-rehearsal and post-rehearsal) nested within rehearsal. Take, note and singer number were also entered as random variables in the models investigating the horizontal tuning across all notes. Take, interval and pair number were inputted as random variables in the models analyzing the major and minor thirds.

A Bonferroni correction was implemented for multiple multilevel linear modelling, and a P value threshold was set at $p = 0.0055$, based on a total of 9 tests (see Table 3).

In order to analyze the verbal interaction between singers during rehearsal in relation to tuning, the total amount of time allocated to each rehearsal was recorded and video recordings of the rehearsal episodes were extracted and uploaded into NVivo (QSR International). The data was transcribed by the fourth author (NP) to produce time-stamped line-by-line verbal utterances of the rehearsal episodes. Episodes of singing were also noted. The length of time allocated to each speech unit or singing episode was recorded in NVivo during the transcription process. Further analysis was performed to identify the points at which singers worked on tuning during their rehearsals. From this data, the amount of time spent on tuning (overall, and by bar/chord), and the nature of the discussion and methods used to address each tuning "problem" were explored.

RESULTS

Horizontal tuning

Visual inspection of the horizontal analysis of tuning clearly demonstrates that each singer was closer to equal temperament than just intonation, and this distinctive behavior was consistent and repeatable during and across rehearsals. This is illustrated in Figure 2 and Figure 3, showing the f_o deviations computed against equal temperament and just intonation for the soprano calculated for each take in rehearsal 1 and rehearsal 5, respectively. The analysis demonstrates that the soprano tended towards equal temperament in both rehearsals and across repetitions within each rehearsal. Complete pitch-drift analysis for each singer/note/take/rehearsal is reported in Appendix A-E. Based on these results, the inferential analysis of tuning during and across rehearsal was based on deviation from equal temperament, rather than just intonation.

Results from the multilevel linear modelling show that, compared with rehearsal 1, the measured deviation from ET was slightly sharper in rehearsal 2 ($\beta = 4.8$, $t(6120) = 3.1$, $P < 0.01$), and flatter in rehearsal 4 ($\beta = -19.8$, $t(6120) = -12.6$, $P < 0.001$) and rehearsal 5 ($\beta = -12.2$, $t(6120) = -7.8$, $P < 0.001$), as shown in Figure 4A. The β – fixed effect coefficients – indicate that for each 1 unit increase in the predictor being considered, the effect of the given predictor changes by the amount specified by the β coefficient. For example, for each 1 unit increase in the tuning of rehearsal 1, tuning computed in rehearsal 2 increased by 4.8 units. Deviation from equal temperament tended to be flatter post-rehearsal in rehearsals 1 to 4 (see Figure 4B), but there was no significant difference pre-rehearsal and post-rehearsal in rehearsal 5. The variance partition coefficient (VPC) among singers and notes was 0.0206 and 0.0248, which demonstrates that only 2% and 2.5% of the variability of tuning can be attributed to singers and notes, respectively. The variability among takes was 16.2%, which indicated that the measured deviation from ET might have changed during repetitions. For

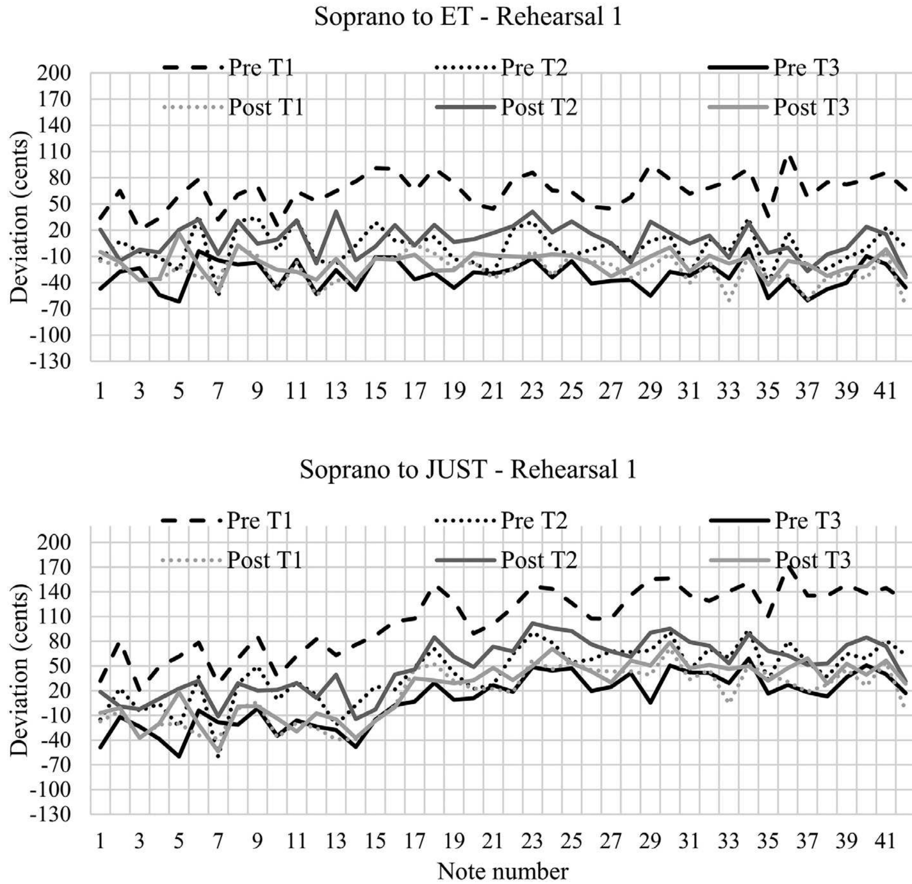


FIGURE 2. Measured deviation from equal temperament (ET, top row) and just intonation (JUST, bottom) of the soprano computed for each note (note 1 to 42), stage (pre-rehearsal and post-rehearsal) and take (T1-T3) during the first rehearsal, R1. Notes are normalized to the first tenor note, F3, which is the tonic of the piece used in the study. Maximum and minimum values on the y-axis have been fixed to allow comparison between the two graphs.

this reason, an ANOVA test was run to test the effect of take. Results show that the take order had a significant effect, $F(2, 6173) = 340.8$, $P < 0.001$, and that the deviation tended to be slightly flatter across repetitions though still closer to ET, as demonstrated by the *post hoc* comparisons using the Bonferroni correction (see Figure 4C).

Results from the multilinear modelling based on the SD of measured deviation from ET show that, compared with the first rehearsal, tuning deviation was more consistent in rehearsal 2 ($\beta = -5.5$, $t(134) = -3.5$, $P < 0.001$), rehearsal 3 ($\beta = -8.1$, $t(134) = -5.2$, $P < 0.001$), rehearsal 4 ($\beta = -5.3$, $t(134) = -3.3$, $P < 0.01$), and rehearsal 5 ($\beta = -4.6$, $t(134) = -2.9$, $P < 0.01$). Tuning was gradually more consistent during the first three rehearsals, but it did not change significantly pre-rehearsal and post-rehearsal, as shown in Figure 5A and Figure 5B. The VRP among takes and singers was 6.2% and 54.5% respectively, suggesting that the consistency of tuning across rehearsals might vary with the ensemble rehearsing. An ANOVA was run to further investigate the role of the singer on the consistency and, as expected, results confirmed a significant effect of singer $t(4, 145) = 29.73$, $P < 0.001$, which was significantly associated with the consistency of singer 5. Tuning of singer 5 was less consistent than that of the other singers, as shown in Figure 5C.

The analysis of the dispersion of tuning across rehearsals shows that the range of tuning deviation from ET was narrower in the third rehearsal compared with the first, $\beta = -33.0$, $t(134) = -3.7$, $P < 0.001$, as shown in Figure 6A. Tuning range did not differ significantly pre-rehearsal and post-rehearsal, as shown in Figure 6B. The variability of the primary effects of rehearsal among take and singers was 8.8% and 38.5% respectively, suggesting that these results might change if different singers were to take part in the study. An ANOVA was conducted to investigate further the effect of singer, and results confirmed that the dispersion differed significantly according to the singer $t(4, 145) = 16.1$, $P < 0.001$. The spread of tuning was wider in singer 5 compared with the other singers, as shown by Bonferroni *post hoc* comparisons (see Figure 6C).

Vertical tuning

The average size of the major thirds was 392.17 cents with a standard deviation of 27.56 cents. This is slightly closer to just intonation (386 cents) than ET (400 cents), and, together with the wide spread, indicates examples of both ET and just intonation, as shown in Figure 7A. The stability of the thirds did not change significantly across rehearsals or pre-rehearsal and post-rehearsal. The variability among interval number,

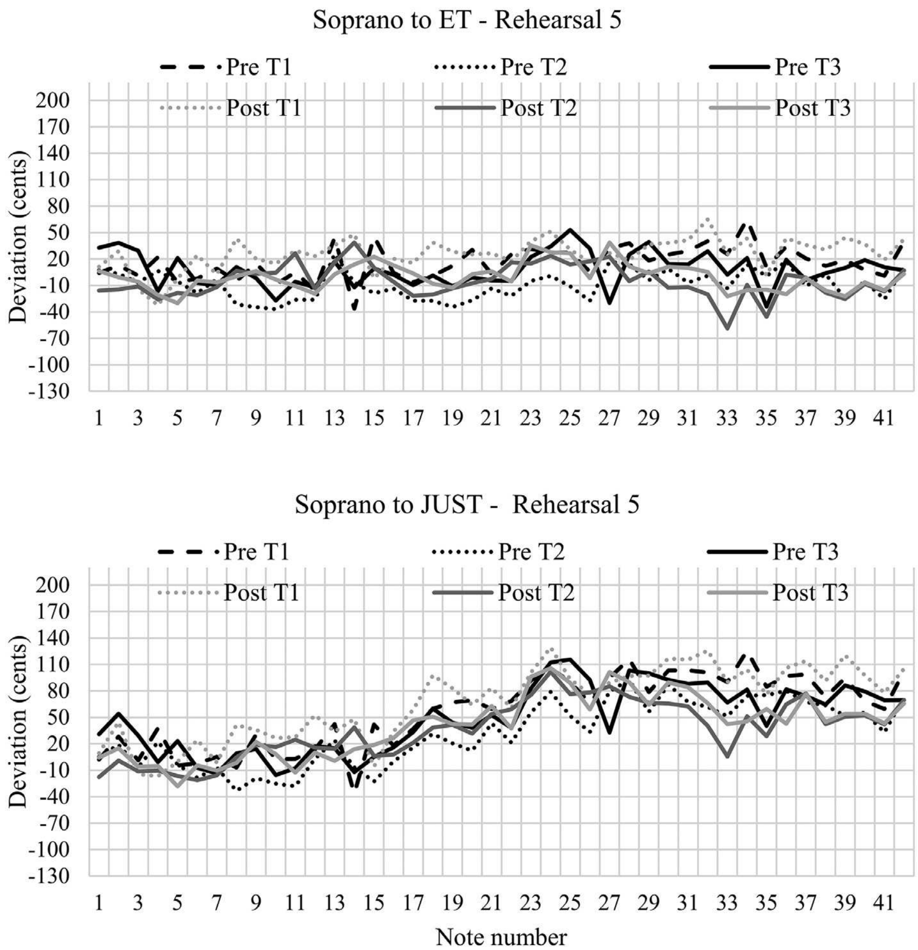


FIGURE 3. Measured deviation from equal temperament (ET, top row) and just intonation (JUST, bottom) of the soprano computed for each note (note 1 to 42), stage (pre-rehearsal and post-rehearsal), and take (T1-T3) during the last rehearsal, R5. Notes are normalized to the first tenor note, F3, which is the tonic of the piece used in the study. Maximum and minimum values on the y-axis have been fixed to allow comparison between the two graphs.

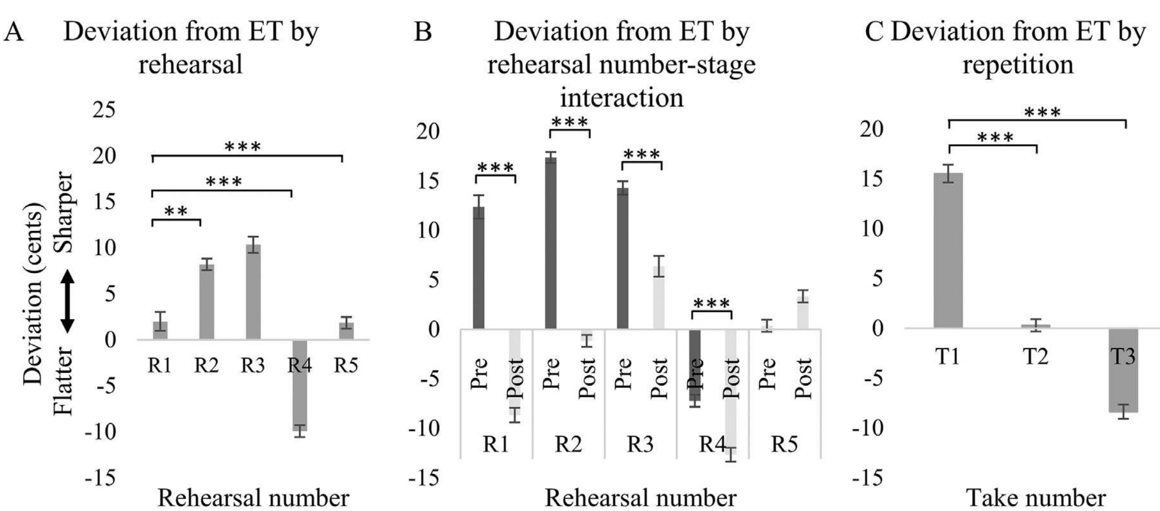


FIGURE 4. Deviation of tuning from equal temperament (ET): A) by rehearsal number (R1-R5); B) by interaction between rehearsal number (R1-R5) and rehearsal stage (pre-rehearsal and post-rehearsal); and, C) by repetition from take 1 to take 3 (T1-T3). Error bars represent 95% CI of *the mean.* ** = $P < 0.01$; *** = $P < 0.001$.

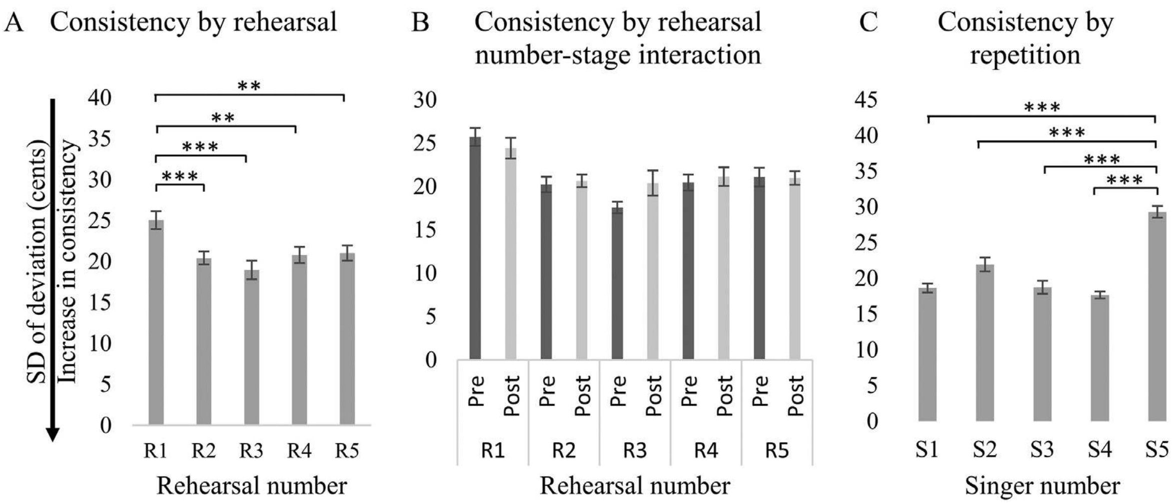


FIGURE 5. Consistency of tuning: A) by rehearsal number (R1-R5); B) by interaction between rehearsal number (R1-R5) and rehearsal stage (pre-rehearsal and post-rehearsal); and, C) by singer (S1-S5). Error bars represent 95% CI of the mean. $^{**}P < 0.01$; $^{***}P < 0.001$.

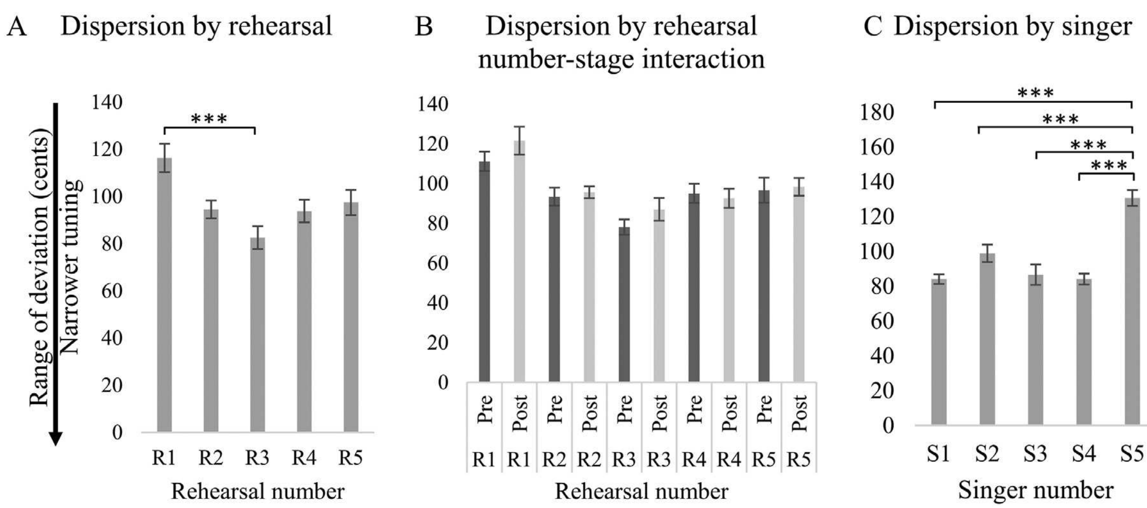


FIGURE 6. Dispersion of tuning: A) by rehearsal number (R1-R5); B) by interaction between rehearsal number (R1-R5) and rehearsal stage (pre-rehearsal and post-rehearsal); and, C) by singer (S1-S5). Error bars represent 95% CI of the mean. $^{***}P < 0.001$.

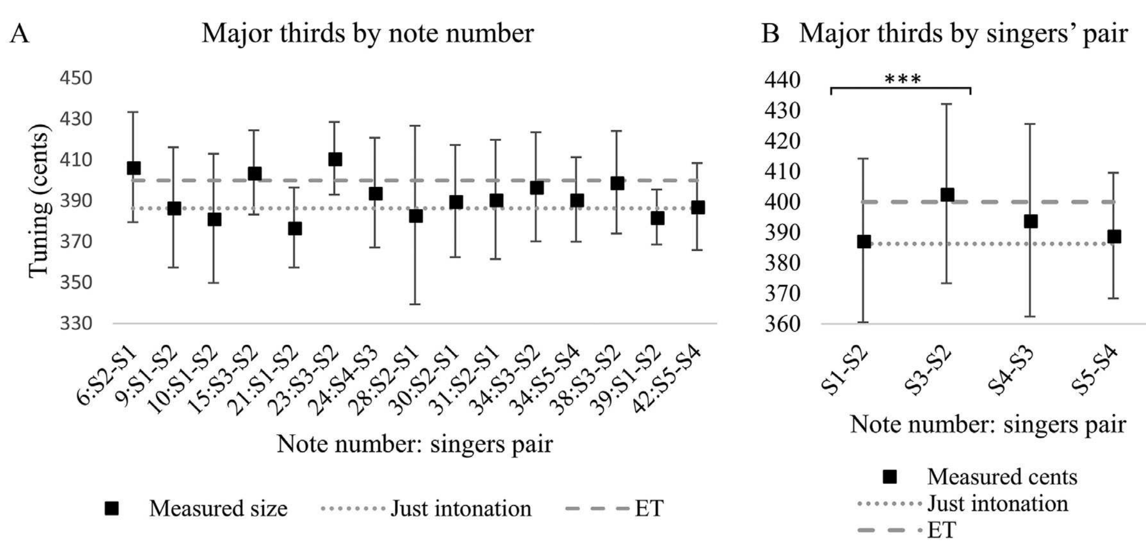


FIGURE 7. Tuning of major thirds: A) by note number, and B) by singers' pair. Error bars represent SD of the mean. $^{***}P < 0.001$.

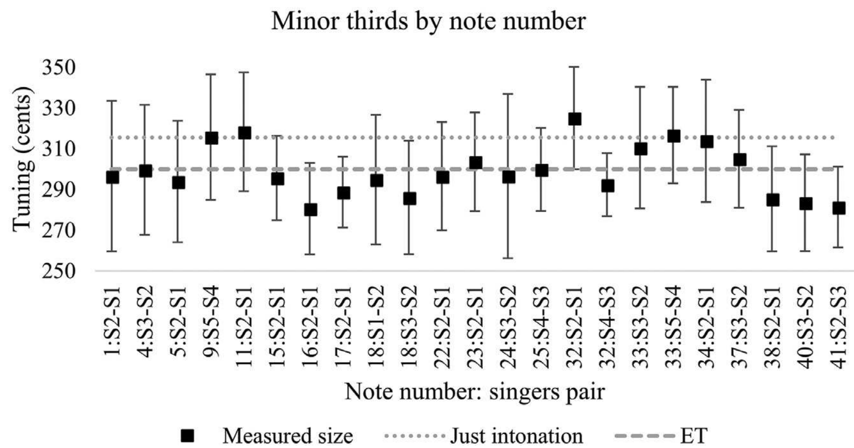


FIGURE 8. Tuning of minor thirds by note number.

pair and take was 5.2%, 5.9%, and less than 0.1%, respectively. Considering the significant effect of singer on the horizontal tuning, an ANOVA was conducted to test whether tuning of the thirds changed according to the singers performing. Results demonstrate a significant effect of the pair of singers, $t(3, 438) = 9.0$, $P < 0.001$, which was associated with the pair S3-S2 and S1-S2, as shown by Bonferroni *post hoc* comparisons (see Figure 7B). Singers 1 and 2 tuned the major thirds closer to just intonation, but singers 2 and 3 tuned closer to ET. Another ANOVA was also conducted to test the effect of note number, and results show the major thirds tuning changed significantly based on the interval considered, $t(14, 427) = 4.1$, $P < 0.001$.

The consistency and range of the major thirds did not differ across rehearsals or pre-rehearsal and post-rehearsal, and the variance partition coefficient among repetitions was 8% in relation to the consistency and 15.2% for the range of major thirds. Further tests were then conducted to investigate the role of take, and ANOVAs show that the consistency and dispersion of tuning of the major thirds did not differ across repetitions.

The average size of the minor thirds was 299.13 cents, with a standard deviation of 29.28 cents, indicating that the tuning of the minor thirds was closer to ET (300 cents), than just intonation (315.6 cents), as shown in Figure 8. The tuning stability, as indexed by the size of interval, did not differ pre-rehearsal and post-rehearsal, or across rehearsals. The variability among minor thirds, pair and take was 15.4%, 1.3% and less than 0.1%, respectively. An ANOVA on the minor thirds number confirmed a significant effect of the interval number on the tuning of the minor thirds, $t(22, 646) = 6.4$, $P < 0.001$. The consistency and range of the minor thirds did not change across the five-rehearsal sessions or pre-rehearsal and post-rehearsal, and the variance partition coefficient among take 1 to 3 was less than 0.1% in relation to both consistency and range.

Verbal interactions during rehearsal

Duration and frequency of verbal utterances relating to tuning

Based on the transcribed verbal utterances, the amount of time dedicated to tuning was summarized as a percentage of the total duration of rehearsal time in each session. Table 1 shows the total rehearsal time and time spent on tuning for the whole study period.

No verbal utterances on tuning topics were observed in rehearsal 2, which was shortened due to one member being indisposed for a time. For this reason, it was not included in the following analyses. Over the entire 5 rehearsal sessions the singers allocated 19% of their rehearsal time to tuning; however, a reduction in time allocated to tuning was apparent across the study period (Figure 9).

From rehearsals 1, 3, 4 and 5, work on specific bars and chords were identified using the verbal interaction data, and is summarized in Figure 10.

Chords on which the group spent most time tuning is reported in Figure 11. Chords of interest were identified as chords 30, 32, 10, 24, and 26, each of which the group dedicated at least 20 seconds of rehearsal time.

Rehearsal strategies for tuning

The verbal interaction data also revealed the strategies used by the group for tuning. These included the identification of problem areas, and proposed ways of dealing with tuning issues. Methods of identification included drawing attention to problem bars, “fuzzy” chords, specific intervals that were

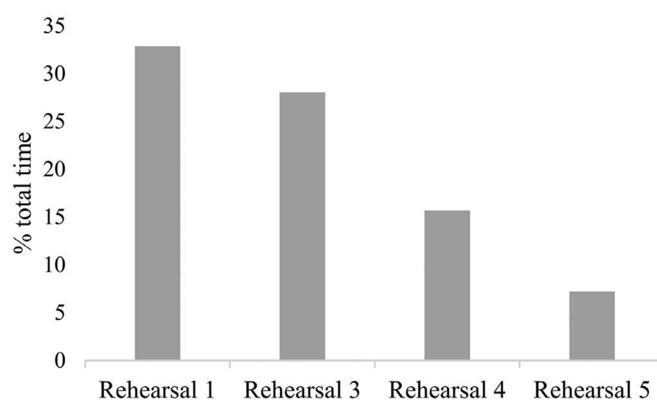
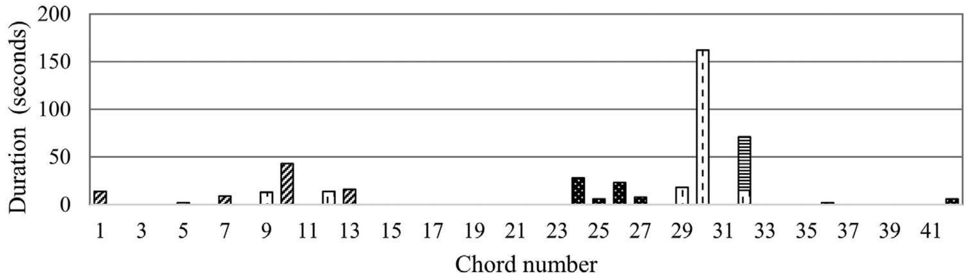


FIGURE 9. Allocation of time to tuning tasks as percentage of total rehearsal time.



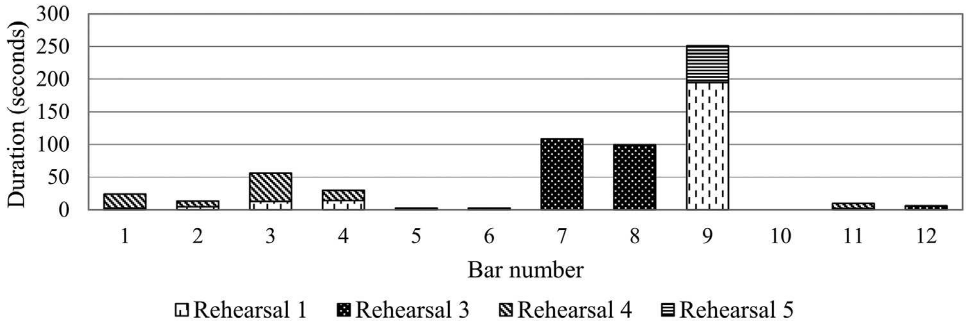


FIGURE 10. Time spent on tuning by bar (top row) and chord (bottom row) for rehearsals 1, 3, 4, and 5.

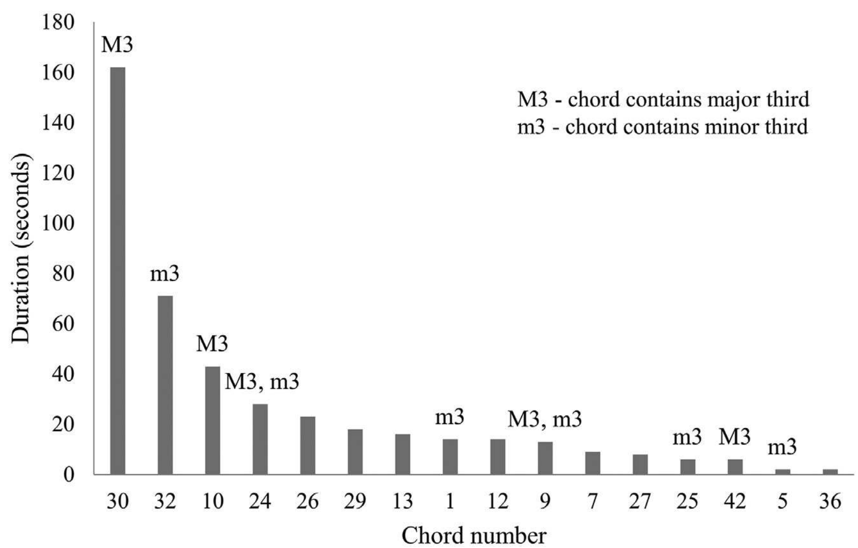


FIGURE 11. Amount of time allocated to tuning, by chord number.

hard to tune, or problem notes in a chord, such as where notes were doubled, or where they created unusual harmonies. A range of strategies were adopted for solving tuning problems as they were identified. These included: i) running or repeating a short section, single bar or chord; ii) separating out parts so that just two or three voices could be isolated; iii) singing a progression more slowly, encouraging each other to listen in a more focused way; and, iv) rebalancing chords so that certain voices could be stronger. In some instances, individuals explicitly stated how they were planning to adjust their tuning within the context of a chord. For example, Singer 3 describes how she is deliberately lowering a minor third, and advises Singer 2 to lower her minor second;

> "I was trying to pull the F down, then", (Singer 3, chords 23, 25 and 27, rehearsal 3)

> "I think it might settle if you really make that semitone close, so you can sit down lower." (Singer 3 to Singer 2, chord 30, rehearsal 1).

In addition, Singer 4 made direct reference to the tuning system he was adopting for a specific chord;

> "I'm going to do an equal tempered third at the end this time" (Singer 4, chord 42, rehearsal 5)

There was no other discussion or direct reference to tuning systems during the remainder of the rehearsals. Aside

TABLE 4.
Tuning Strategies for Minor and Major Thirds

Tuning Strategies	Minor Thirds Chord Number (Voices)	Major Thirds Chord Number (Voices)
Tuning "doubled" notes	9 (S5, S4)	9 (S1, S2)
	24 (S3, S2)	10 (S1, S2)
	25 (S4, S3)	24 (S3, S4)
	32 (S4, S3)	42 (S4, S3/S2)
Balancing voices	5 (S2, S1)	30 (S1, S2)
	32 (S2, S1)	
Tuning of whole chord	1 (S2, S1)	9 (S1, S2)
	32 (S4, S3)	10 (S1, S2)
		30 (S1, S2)
Tuning melodic intervals	1 (to 2) (S2)	
	24 and 25 (S3, S5)	
Aiming for equal temperament		42 (S4, S5)

from this one example where a tuning system was explicitly mentioned, tuning strategies for chords containing thirds were primarily focused on balance, matching intonation of specific pitches in unison or octaves, or tuning a whole chord which included thirds (see Table 4). However, there were some examples where singers referred to tuning of specific melodic or harmonic thirds. Singer 2 expresses difficulty with tuning of melodic major thirds in bars 1 and 11:

> "I'm very conscious of my falling major thirds in bar 1 and bar 11, I'm finding them quite hard to tune, I don't know why." (Singer 2, rehearsal 4, chords 1 to 4 and 36 to 39).

In rehearsal 3, singer 4 draws attention to a harmonic minor third, which prompts singers 3 and 5 to work on tuning of chord 24:

> "I think I'm hearing the minor third between the F and the D, [singer 3 and singer 5] as too wide...in other words the D on the third beat of the bar as too low." (Singer 4, chord 24, rehearsal 3)

DISCUSSION

This study investigated the evolution of tuning across five rehearsal sessions in a newly formed, semi-professional singing quintet during a first term of study. The analysis of tuning was based on a mixed method that combined the physical measurement of tuning and the investigation of the verbal interactions between singers during rehearsals. The physical measurement was based on the analysis of the f_o deviation computed against the expected equal temperament and just intonation, and measured horizontally (ie, for each note/take/singer/rehearsal) and vertically (ie, in relation to the major and minor thirds of the piece). The verbal interactions between singers were investigated through time-stamped transcriptions of rehearsal discussions, to identify verbal interactions related to tuning, the tuning strategies adopted, and their location (bar or chord) within the piece.

Each singer in this study consistently tended towards equal temperament during and across rehearsals, as demonstrated by the pitch drift analysis. These results corroborate previous investigations conducted by Devaney et al[4], showing no evidence of pitch drift in a Benedetti's three-part exercise, and contrast the findings from Howard,[5,6] observing pitch drift in a four-part piece, in which the chords were linked via a tied note in each case, composed for the study. These findings suggest that tuning in singing ensemble might depend on the specific melodic/harmonic characteristics of the piece performed as well as the individual singers and combination of singers performing.

Furthermore, compared with the first rehearsal, intonation computed against ET was significantly flatter in rehearsal 4 and 5, ie towards the end of the first term of study. It was also flatter with repeated performances, and, in most rehearsals, post-rehearsal. Tuning deviation from ET was less consistent in the first rehearsal, compared with the rest of the rehearsals, as shown by the SD of the measured deviation. This is not surprising, as the singers did not know each other before the first rehearsal, and did not practice the piece before. The consistency did not change pre-rehearsal and post-rehearsal, but gradually improved in the first three rehearsals. The tuning of singer 5 was significantly less consistent and wider compared with the other singers, as quantified by the SD and range of measured deviation, respectively, but still highly accurate.

Tuning deviation from ET was more consistent and narrower in the third rehearsal, which was the anticipated midpoint of the first term of study, although, due to some last-minute issues, the final rehearsal date was moved, and consequently the anticipated and actual midpoint were different. Therefore, rehearsal 3 was at the time the anticipated midpoint, and played a crucial role in the consistency and dispersion of tuning. The role of this rehearsal is consistent with the group development theory advanced by Gersick,[33] suggesting a turning point in the development, halfway through the process of working towards a shared goal, in which there is a transition from exploration mode to action planning mode.

The size of the major thirds was slightly closer to just intonation, with examples of both just and ET system across the piece. This did not change within (ie, pre-rehearsal and post-rehearsal) and across rehearsals (R1-R5), and repetitions (T1-T3). Chord number and pair, however, did significantly affect the size of the major thirds. The pair singer 1 and 2 (S1-S2) tuned their major thirds closer to just intonation, but the pair singer 2 and 3 (S2-S3) closer to ET. These results suggest that the tuning of major thirds might change with singer and the harmonic/melodic content of the piece. The highly variable size of the major thirds across pair of singers measured in this study also corroborates the results from perception studies, showing different subjective preferences for the size of major third dyads.[11]

Intonation of the minor thirds was clearly closer to ET; this did not change across repetitions and rehearsals, or pre-rehearsal and post-rehearsal. Chord number significantly affected the tuning, suggesting, similar to the finding with the major thirds, that intonation of the minor thirds might be context-specific. The variability of the minor and major thirds based on the chord number, including examples of intonation close to both ET and just intonation, is in line with previous results found by Devaney et al[4] when investigating minor and major thirds in a three-part progression written by Benedetti. The consistency and distribution of the major and minor thirds' tuning, as quantified by the SD and range of the thirds' size respectively, did not differ pre-rehearsal and post-rehearsal, and across rehearsals and repetitions, suggesting that this tuning behavior was highly repeatable in relation to the minor and major thirds and typified the ensemble.

The analysis of the verbal interactions in relation to aspects of tuning that emerge spontaneously during rehearsal demonstrates that singers allocated 19% of the total time rehearsing to aspects of tuning. This indicates that tuning was a consistent feature of rehearsal in this ensemble, in line with previous research conducted among professional *a cappella* vocal ensembles.[18] The time spent on tuning decreased during the study period, and this might be understood in light of previous investigations showing a shift from "basic" tasks, such as work on intonation, in early rehearsals to "interpretative" tasks, such as expressive intentions, in later rehearsals.[24] The ensemble made use of a range of strategies to improve the overall tuning, such as repeating a short section, single bar, chord, tuning "doubled" notes, and isolating and rebalancing two or three voices, so certain voices could be stronger. These strategies were also applied specifically to chords containing major and minor thirds, and indeed a number of these chords (eg, chords 30, 32, 10, and 24) were the focus of the most time in solving tuning issues. Most tuning time was allocated to major thirds, but this did not appear to be deliberate strategy on their part, as there was little explicit discussion of chord type. The reasons for this are not clear, although it may be that these chords had characteristics that meant the tuning issues were easier to detect, for example, with the presence of doubled notes or octaves. In general, the verbal interactions revealed that methods used for tuning thirds was rather indirect, as the group found ways to identify problems and generally improve the tuning of these chords, rather than focusing on tuning of harmonic intervals or aiming explicitly for adherence to equal temperament or just intonation tuning systems.

The combination of measured intonation horizontally and for vertical thirds with analysis of the verbal interaction reveals a complex picture of the tuning strategies of this quintet. The increased consistency in horizontal tuning with rehearsal, peaking in rehearsal 3, is an expected result, in that as the singers practice they establish their tuning of the piece. This is most probably related to other performance goals including blend and expression, and reflects the findings of the verbal interaction data that they spend less time discussing tuning throughout the term. That the increased consistency is in parallel with overall flatter horizontal intonation is unexpected but suggests that the group are satisfied with these intonation outcomes, also implied by the reduced discussion on tuning in the later sessions. This ensemble overwhelming tune to ET horizontally, avoiding pitch drift, and also present features of just intonation occurring frequently in the tuning of major thirds. The varied but repeated tuning of thirds suggests either context and/or singer specific practice, however this seems to be a spontaneously emerging characteristic based on the absence of specific work to tune thirds, rather the group worked to tune an overall chord or match octaves. A notable increase in time spent on tuning of chords with major thirds is in line with experience reported in the literature that thirds are difficult to tune,[34] but without explicit reference to the thirds within those discussions this data supports the hypothesis that tuning is highly context driven based on a complex number of factors rather than a simple aim to tune thirds within a specific system.

LIMITATION AND FUTURE WORKS

This study was based on a single ensemble performing a five-part piece featuring a simple harmonic context and a homophonic structure. Future investigations should use other ensembles and pieces with greater tonal and rhythmical complexity, to test the replicability of the above findings in more musically complex pieces. The stimulus used in this study was a Bach chorale arranged to facilitate and maximize the accuracy in the analysis of tuning by limiting semitone progressions in any part, but still challenging the singers during the five rehearsal sessions across a first-term of study. The resulting stimulus was a medium-length piece without repeated notes and with few semitones. Parallel octaves and fifths are present in the arrangement, which were strictly forbidden in Bach's harmonies but not uncommon to other kinds of music, such as jazz, popular and folk music. Further investigations could analyze the approach to tuning in an untouched Bach composition, addressing the effect of parallel octaves and fifths on tuning using an untouched Bach's composition in addition to a manipulated piece with some consecutive octaves and fifths.

Singers performed the pieces to the vowel /i/ for consistency with previous investigations analyzing repeated performances in singing ensembles.[26,28,29] It is common to observe slight sharpening in ascending passages when the text makes use of this vowel, although this was not the case in the present study. It would be of interest for further studies to investigate the effect of the chosen vowel on tuning, through repeated performances of a shorter piece sung each time to a different vowel.

This study focused on five lab-based rehearsal sessions, representing five snapshots captured across a first term of study. The ensemble continued to rehearse outside of the study in order to work on other pieces, and was coached between lab-based rehearsals; these extra events were not considered in this study but will have contributed to the development of the group's performance traits. Further research should analyze all rehearsal and coached sessions of a specific and shorter study period to investigate the development of tuning with controlled practice, and in relation to coaching.

The intent of the current investigation was to analyze aspects of tuning that emerged spontaneously with practice. For this reason, singers were not aware of the scope of this study, but were asked to work on producing an expressive performance. Another avenue of research should consider the evolution of tuning when singers are explicitly asked to master tuning. This may shed some more light on the rehearsal strategies that singers consciously apply to excel tuning during singing ensemble performances, as well as determine which tuning systems they are aspiring to.

Finally, singers were invited to master the expressive performance of the piece across rehearsals, pretending that they would have had to perform the piece in form of a concert. Further studies should consider a more realistic situation, with an ensemble working on a piece that will be then also performed on stage.

CONCLUSIONS

This study investigated the evolution of tuning in a newly-formed advanced singing quintet during five rehearsal sessions across a first term of study. Each singer performed closer to ET, avoiding pitch drift based on just intonation predictions, during and between rehearsals and across repetitions. Deviation from equal temperament was flatter towards the ending of the first term compared with the initial rehearsal, and was more consistent and narrower in the third rehearsal. The ensemble tuned the major thirds slightly closer to just intonation, and the minor thirds closer to equal temperament, and these results were consistent within and between rehearsals, but changed based on chord number and singers' pair.

Tuning was an important dimension of rehearsal in this ensemble with 19% of the total time of rehearsal dedicated to tuning, and which showed a decrease over time from rehearsals 1 (33%) to 5 (7%). Singers adopted a range of strategies to solve tuning related issues, including tuning doubled notes, whole chords, specific melodic intervals, and balancing voices.

The above findings contribute to understanding the developmental aspects of tuning in advanced singing ensemble. This study provides an evidenced base and context from which choir directors and coaches can develop their rehearsal strategies and performance goals.

Acknowledgments

The authors would like to thank the singers who took part in the study, and the White Rose College of the Arts and Humanities (WRoCAH) for funding this investigation.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.jvoice.2018.07.015.

APPENDIX A

Measured tuning deviation from expected equal temperament (on the left-hand side) and predicted just intonation (right-hand side) for each singer and repetition across rehearsal 1.

APPENDIX B

Measured tuning deviation from expected equal temperament (on the left-hand side) and predicted just intonation (right-hand side) for each singer and repetition across rehearsal 2.

APPENDIX C

Measured tuning deviation from expected equal temperament (on the left-hand side) and predicted just intonation (right-hand side) for each singer and repetition across rehearsal 3.

APPENDIX D

Measured tuning deviation from expected equal temperament (on the left-hand side) and predicted just intonation (right-hand side) for each singer and repetition across rehearsal 4.

APPENDIX E

Measured tuning deviation from expected equal temperament (on the left-hand side) and predicted just intonation (right-hand side) for each singer and repetition across rehearsal 5.

REFERENCES

1. Hansen C. Tuning the choir. *Music Educ J*. 1964;51:85–89.
2. Havrøy F. ‘You cannot just say: “I am singing the right note”’. Discussing intonation issues with Neue Vocalsolisten Stuttgart, *Music Pract* 2013. Available at: http://www.musicandpractice.org/volume-1/intonation-neue-vocalsolisten-stuttgart/. Accessed March 21, 2015.
3. Ternström S. Choir acoustics—an overview of scientific research published to date. *Int J Res Choral Sing*. 2003;43:1–8.
4. Devaney J, Mandel MI, Ellis DP, et al. Automatically extracting performance data from recordings of trained singers. *Psychomusicol Music Mind Brain*. 2011;21:108–136.
5. Howard DM. Intonation drift in *a capella* soprano, alto, tenor, bass quartet singing with key modulation. *J Voice*. 2007;21:300–315.
6. Howard DM. Equal or non-equal temperament in a capella SATB singing. *Logoped Phoniatr Vocol*. 2007;32:87–94.
7. Howard DM, Brereton J, Welch GF, et al. Are real-time displays of benefit in the singing studio? An exploratory study. *J Voice*. 2007;21:20–34.

8. Devaney J, Ellis DP. An empirical approach to studying intonation tendencies in polyphonic vocal performances. *J Interdiscip Music Stud*. 2008;2:141–156.
9. Potter J. Ensemble singing. In: Potter J, ed. *The Cambridge Companion to Singing*. Cambridge University Press; 2000:158–164.
10. Mayer F. The relationship of blend and intonation in the choral art. *Music Educ J*. 1964;51:109–110.
11. Ternström S, Nordmark J. Intonation preferences for major thirds with non-beating ensemble sounds. In: *Nordic Acoustical Meeting*, Helsinki, 12−14 June 1996-; 1996:359−365.
12. Terasawa H. Pitch drift in choral music. *Music 221A final paper; Center for Computer Research in Music and Acoustic*. CA: at Stanford University; 2004. Available at: https://ccrma.stanford.edu/∼hiroko/pitchdrift/paper221A.pdf. Accessed June 5, 2014.
13. Planchart A. On singing and the vocal ensemble I. In: Kite-Powell J, ed. *The Performer's Guide to Renaissance Music*. 2nd ed. Bloomington, IN, USA: Indiana University Press; 2007:28–41.
14. Blachly A. On singing and the vocal ensemble II. In: Kite-Powell J, ed. *The Performer's Guide to Renaissance Music*. 2nd ed. Bloomington, IN, USA: Indiana University Press; 2007:14–27.
15. Murray Barbour J. Just intonation confuted. *Music Lett*. 1938;1:48–60.
16. Timm E. Intonation facts. *Music Educ J*. 1943;29:19–20.
17. Duffin R. Tuning and temperament. In: Kite-Powell J, ed. *The Performer's Guide to Renaissance Music*. 2nd ed. Bloomington, IN, USA: Indiana University Press; 2007:279–298.
18. Havrøy F. *Alone Together. Vocal Ensemble Practice Seen through the Lens of Neue Vocalsolisten Stuttgart*. Norges musikkhøgskole; 2015.
19. Blank M, Davidson J. An exploration of the effects of musical and social factors in piano duo collaborations. *Psychol Music*. 2007;35:231–248.
20. Davidson J, King EC. Strategies for ensemble practice. In: A Williamon (Ed), *Musical Excellence: Strategies and Techniques to Enhance Performance*. Oxford, UK: Oxford University Press; 2004:105–122.
21. Goodman EC. Ensemble Performance. *Musical Performance: A Guide to Understanding*. Cambridge, UK: Cambridge University Press; 2002:153–167.
22. Williamon A, Davidson JW. Exploring co-performer communication. *Music Sci*. 2002;6:53–72.
23. Chaffin R, Imreh G. Practicing perfection: Piano performance as expert memory. *Psychol Sci*. 2002;13:342–349.
24. Ginsborg J, Chaffin R, Nicholson G. Shared performance cues in singing and conducting: A content analysis of talk during practice. *Psychol Music*. 2006;34:167–194.
25. Ginsborg J, King E. Rehearsal talk: familiarity and expertise in singer-pianist duos. *Music Sci*. 2012;16:148–167.
26. D'Amario S, Daffern H, Bailes F. A longitudinal study investigating synchronization in a singing quintet. *J Voice* 2018. http://dx.doi.org/10.1016/j.jvoice.2018.06.011.
27. D'Amario S, Daffern H. Using electrolaryngography and electroglottography to assess the singing voice: a systematic review. *Psychomusicology Music Mind Brain*. 2017;27:229–243.
28. D'Amario S, Daffern H, Bailes F. A new method of onset and offset detection in ensemble singing. *Logop Phoniatr Vocology* 2018. http://dx.doi.org/10.1080/14015439.2018.1452977.
29. D'Amario S, Daffern H, Bailes F. Synchronization in singing duo performances: The roles of visual contact and leadership instruction. *Front Psychol* ***9***, 2018. http://dx.doi.org/10.3389/fpsyg.2018.01208.
30. Daffern H. Blend in singing ensemble performance: vibrato production in a vocal quartet. *J Voice*. 2017;31:385.e23–385.e29.
31. P Boersma, Praat: a system for doing phonetics by computer. *Glot Int 5,* 2001, 341−347.
32. Howard DM, Angus JAS. *Acoustics and Psychoacoustics*. 5th ed. Oxford, UK: Focal Press; 2017.
33. Gersick CJG. Time and transition in work teams: toward a new model of group development. *Acad Manag J*. 1988;31:9–41.
34. Covey-Crump R. Vocal consort style and tuning. In: Paynter J, ed. *Companion to Contemporary Musical Thought, 2*. London and New York: Routledge Reference; 1992:1020–1050.