Voice Quality of Choir Singers and the Effect of a Performance on the Voice

Iris Meerschman [,*], Evelien D'haeseleer [2], Heleen Cammu [3], Imke Kissel [3], Tine Papeleu [3], Clara Leyns [3], Julie Daelman [3], Jelena Dannhauer [4], Laura Vanden Abeele [5], Valerie Konings [5], Erik Demarbaix [5], Kristiane Van Lierde [6]

[1] Centre for Speech and Language Sciences, Department of Rehabilitation Sciences, Ghent University, Ghent, Belgium. Electronic address: iris.meerschman@ugent.be.

[2] Centre for Speech and Language Sciences, Department of Rehabilitation Sciences, Ghent University, Ghent, Belgium; Royal Conservatory Brussels, Musical Department, Brussels, Belgium.

[3] Centre for Speech and Language Sciences, Department of Rehabilitation Sciences, Ghent University, Ghent, Belgium.

[4] World Choir Games 2021, Interkultur, Fernwald, Germany.

[5] World Choir Games 2021, Koor & Stem, Antwerpen, Belgium.

[6] Centre for Speech and Language Sciences, Department of Rehabilitation Sciences, Ghent University, Ghent, Belgium; Faculty of Humanities, Department of Speech-Language Pathology and Audiology, University of Pretoria, Pretoria, South Africa.

*Address correspondence and reprint requests to Iris Meerschman, Centre for Speech and Language Sciences, Department of Rehabilitation Sciences, Ghent University, Campus UZ Gent, Corneel Heymanslaan 10, 2P1, entrance 69, 9000 Ghent, Belgium. Email: iris.meerschman@ugent.be

Summary

Background: The voice use of choir singers is understudied despite the imbalance of high vocal demands versus low vocal education, and consequently increased risk for voice problems. Also, there is a lack of information on the effects of a performance on choristers' voices. Available studies included performances of at least one hour. To date, no studies investigated the effects of a choir performance with a duration resembling vocal warm-ups.

Purpose: The first purpose of this study was to determine the voice quality, capacities, symptoms and voice-related quality of life of choir singers. Secondly, the effect of a short choir performance, resembling warm-up duration (15 minutes), on the choristers' voices was investigated.

Methods: A randomized controlled trial was used. Thirty adult choir singers (25 women, 5 men; mean age: 32 years) were assigned randomly to an experimental group or a control group. Participants in the experimental group sung in choir for 15 minutes immediately after their pre voice assessment, whereas the control group was instructed to have standard voice use (one-on-one conversation with the investigator, no singing) across that time span. A second voice assessment was repeated afterwards.

Results: The choir singers showed excellent voice quality and capacities with mean scores on the Dysphonia Severity Index and Acoustic Voice Quality Index of 7.5 and 2.0, respectively.

Auditory-perceptually, the mean grade score was 5/100 corresponding with a normal to mildly deviant voice quality. Patient-reported outcome measures showed mean deviant scores, indicating a considerable singing voice handicap. The choir singers seem vulnerable for stress with a high occurrence rate of 76.7% (23/30). Compared with the control group, the Dysphonia Severity Index significantly improved, whereas the self-perceived presence of vocal fatigue and complaints increased after 15 minutes of choir singing. Fundamental frequency increased in both groups, being more outspoken in the experimental group.

Conclusions: Choir singers show excellent voice quality and capacities, that further improve after a short choir performance of 15 minutes. Vocal fatigue and complaints, on the other hand, already increased after that short time span. Realizing that vocal load is much higher in real-life rehearsals, competitions and performances, choristers deserve and need a qualitative voice training and a strict follow-up. Future research should focus on effective vocal warm-up and cool-down programs for this population.

Keywords: Voice; Voice quality; Choir singers; Performance; Effect; Randomized controlled trial

INTRODUCTION

Choral singing requires extra vocal demands in an attempt to match the singing group.[1] Compared to solo singing, the focus is on contribution and blending rather than being clearly heard, which asks additional flexibility and control of the respiratory and phonatory mechanisms.[1,2]

Despite these high vocal demands, choir singers often did not receive formal singing training and have limited vocal knowledge.[2, 3, 4, 5, 6] Consequently, amateur choristers are at higher risk for developing voice disorders, especially female singers and those with a vocally demanding occupation.[2,5] The prevalence of voice disorders in choir singers seems similar to that of teacher students and teachers.[2] They frequently report vocal symptoms such as throat clearing, globus sensation or pain, a strained or tired voice, and hoarseness.[2,6] A recent study of Robotti et al.[7] showed higher self-reported voice and vocal tract impairments in amateur choir singers compared to control subjects. Nevertheless, both professional and amateur singers seem concerned about maintaining a healthy vocal mechanism and are interested in expanding their knowledge on vocal hygiene and function.[3]

Although choristers generally have minimal or no formal singing training, they still show superior voice quality and capacities compared to non-singers.[1,4,8,9] More specifically, lower shimmer and noise-to-harmonic ratio, and higher smoothened cepstral peak prominence were found in this population.[1,4] Also, choir singers show higher (slow) vital capacity and a wider vocal range.[1,8,9] Lastly, excellent scores on the Dysphonia Severity Index, with medians ranging from 5 to 11 (depending on the age), have been reported.[10]

To fully and efficiently engage these vocal capacities and mitigate phonotrauma, warming-up the voice before a choral performance is indispensable.[11, 12, 13] Vocal warm-up prepares the phonatory apparatus for the requirements of singing.[14,15] Warm-up exercises are hypothesized to increase the temperature and blood flow of vocal musculature which reduces its viscosity and resistance.[15-17] This might consequently lower the phonation threshold pressure and therefore lead to more economic voice use.[16] Preferred duration of vocal warm-up is estimated between 5 and 15 minutes.[15,18, 19, 20] Nonetheless the proven advantages, there seems to be a general lack of vocal warm-up in the majority of choir singers.[6,21] If they do warm-up, then it is mostly limited to singing musical notes or a song before the performance,[1,21] instead of using more specific warm-ups, such as semi-occluded vocal tract (SOVT) exercises.

Therefore, it is debatable whether choral singing itself can be seen as a warm-up or a fatiguing activity.[22] In the first case, choral singing would promote the individual's voice quality and capacities. If fatigue occurs, on the other hand, these capacities may deteriorate and signs of discomfort might take over. Susceptibility to vocal fatigue depends on several factors, including the experience, professionalism, and musical style of the singers,[13] and it is individual specific.[23] Besides these factors, also the duration of singing will be of impact.[15]

To date, research on the effects of a singing performance on choristers' voices is scarce. Kitch et al.[22] reported a deterioration in terms of voice range and perturbation measures in tenors after 90 minute of choir singing. Onofre et al.[13] found a significantly increased fundamental (f_o), highest and lowest frequency, and a decreased peak amplitude variation after 60 minute of choir singing in sopranos. Also, self-perceived symptoms, such as a dry throat, hoarseness and fatigue increased after singing for one hour. In an earlier study of the same authors,[15] f_o increased and jitter decreased after 60 minutes of soprano choir singing. To date, no randomized controlled trials investigated this matter.

In summary, the voice use of choir singers is understudied,[7] despite the imbalance of high vocal demands versus low vocal education, and consequently increased risk for voice problems. Therefore, the first objective of this study was to determine the voice quality, capacities, symptoms and voice-related quality of life of choir singers. Based on the available literature, it can be hypothesized that the choir singers will show excellent voice quality and capacities,[1,4,8, 9, 10] although symptoms might be frequent and the voice related quality of life might be considerably low.[2, 3, 4, 5, 6, 7] Secondly, there is a lack of information on the effects of a performance on choristers' voices. Available studies included performances of at least one hour. To date, no studies investigated the effects of a choir performance with a duration resembling vocal warm-ups. Therefore, the second objective of this study was to investigate the effect of a short choir performance (15 minutes) on the choristers' voices, using a randomized controlled trial. It is possible that the voice quality and capacities improve after the performance (as an effect of vocal warm-up) or that vocal discomfort and complaints increase (as a sign of vocal fatigue).[15,22]

METHODS

This study was approved by the Ethics Committee of Ghent University Hospital (BC-10476).

Participants

Data were recruited at the 11th edition of the World Choir Games 2021 in Antwerp, Belgium. This is the world largest international choir competition that welcomed more than 300 non-professional choirs form more than 50 nations. An information flyer was shared before the event and choir singers who were interested to participate were scheduled for the study. Eventually, thirty adult choir singers, 25 women and 5 men, with a mean age of 32 years (SD: 11.2, range: 21 – 67 years) participated. Most of them (*n* = 26, 86.7%) were Belgian, two were Dutch (6.7%), one was German (3.3%) and one was German-American (3.3%).

Twenty-three participants were working (76.7%) with 30% of them being professional voice users, one was retired (3.3%) and six were students (20%).

Twenty-eight (93.3%) of the participants were amateur choristers who received no income from singing. One participant was a semi-professional singer (second source of income besides singing) and one was a professional singer (singing as the main source of income). Just over half of them (*n* = 17, 56.7%) received singing training in terms of private lessons and/or music education at a conservatory. Mezzo-sopranos represented the largest part of the singers (*n* = 10, 33.3%), followed by altos (*n* = 8, 26.7%), sopranos (*n* = 7, 23.3%), baritones (*n* = 2, 10%), bassos (*n* = 2, 10%) and tenors (*n* = 1, 3.3%). The song genres varied from musical to jazz to rock, but the most common were pop and (contemporary) classical.

Data collection and design

A randomized controlled trial was used. Participants were assigned randomly to an experimental group (*n* = 15; 12 women, three men) or a control group (*n* = 15; 13 women, two men). Participants in the experimental group sung in choir (with a subgroup, no complete choir) for 15 minutes immediately after their pre voice assessment, whereas the control group was instructed to have standard voice use (one-on-one conversation with the investigator, no singing) across that time span. A second voice assessment was repeated afterwards.

Musical genres during the 15 minutes of choir singing in the experimental group were Contemporary Commercial Music and Contemporary Classical Music. The repertoire was familiar, all choirs performed previously rehearsed work. The singers were standing, organized by vocal part, and they were all in their familiar choral situation, however, without the conductor and in a subgroup of the choir.

The conversations in the control group were not standardized. The investigator asked questions out of interest and the participant was therefore talking most of the time. There were no opportunities to hydrate or rest during the 15 minutes in both conditions.

Voice assessment

Questionnaire and patient-reported outcome measures (PROMs)

At the pre assessment, patients filled in a questionnaire gathering demographic and voice-related data (e.g. singing experience, vocal load, (history) of voice disorders, influencing factors). Also, the Singing Voice Handicap Index 10 (SVHI-10), the Vocal Tract Discomfort Scale (VTDS) and the Vocal Fatigue Index (VFI) were administered. At last, a visual analogue scale (VAS) was completed determining current vocal fatigue and complaints. The same VAS was repeated at the post assessment to determine any short-term changes.

SVHI-10. The shortened version of the SVHI[24,25] was used to measure the impact of singing voice problems. The index includes six physical, two functional and two emotional statements that need scoring on a 5-point Likert scale (0-never, 1-almost never, 2-sometimes, 3-almost always, 4-always). The total score is the sum of item scores with a maximum of 40. A higher index indicates a greater perceived singing voice handicap.

VTDS. The VTDS[26,27] consists of eight sensations that can be felt in or around the throat: burning, tight, dry, aching, tickling, sore, irritable, and globus. Each item needs to be scored on frequency (0-never, 1-seldom, 2-sometimes, 3-more than sometimes, 4-often, 5-very often, 6-always) and severity (0-no, 1-almost no, 2-limited, 3-more than limited, 4-moderate, 5-more than moderate, 6-severe perception) using a 7-point Likert scale. The total VTDS score (sum of frequency and severity) can range from 0 to 96 with higher scores indicating more perceived discomfort.

VFI. The VFI[28] identifies vocal fatigue by analyzing three factors: (1) tiredness of voice and voice avoidance; (2) physical discomfort associated with voicing; and (3) improvement of symptoms with rest. It is a 19-item index using a 5-point Likert scale (0-never, 1-almost never, 2-sometimes, 3-almost always, 4-always). The total score can range from 0 to 76, with higher scores indicating more perceived vocal fatigue.

VAS vocal fatigue and complaints. Participants were asked to quantify ten statements related to the current presence of vocal fatigue and complaints on a 100 mm VAS with 0 mm (left) indicating completely disagree and 100 mm (right) indicating completely agree. The statements were as follows: (1) My voice currently feels tired when I talk, (2) Currently, I am experiencing a painful throat during voice use, (3) I currently experience effort when talking, (4) It currently feels like 'work' when I use my voice, (5) My voice currently has insufficient power, (6) My voice is weak at the moment, (7) Currently, it is difficult to project my voice, (8) I currently have a hard time singing high notes, (9) It is difficult to sing quietly right now, (10) My voice currently sounds hoarse. The total score on the VAS (maximum of 1000) was used for further analysis.

Voice quality and capacities

At both the pre and post assessment, voice quality and capacities were determined with instrumental (maximum phonation time, acoustic analysis, multiparametric indices) and subjective (auditory-perceptual) evaluations.

Maximum phonation time. Maximum phonation time (MPT) was determined by asking the participants to sustain the vowel /a:/ at their habitual pitch and loudness after a maximal inspiration in free field while standing. The production was modeled by the assessor, and the participants received verbal encouragements to produce the longest possible sample. MPT was measured with a chronometer, and the best trial of three attempts was retained for further analysis.

Acoustic analysis. All voice samples were recorded with an AKG Lyra USB condenser microphone and the software program Praat (v.6.2[29]). The microphone-to-mouth distance was 15 cm and the participants were standing during the assessment. Fundamental frequency (f_o) and jitter were analyzed based on a midvowel segment of /a:/ (3s) produced at habitual pitch and loudness, following an automatic series (counting to three). Participants also read the phonetically balanced text "Papa en Marloes",[30] "Der Nordwind und die Sonne"[31] or "The Rainbow Passage"[32] to determine the Acoustic Voice Quality Index and its acoustic parameters (see below: multiparametric indices).

Afterwards, the extremes of the voice range profile (lowest and highest frequency [F-low, F-high] and lowest and highest intensity [I-low, I-high]) were established by asking the participants to produce the vowel /a:/ at their minimal pitch, minimal intensity, maximal pitch, and maximal intensity, respectively. For each extreme, the participants had three attempts and the best trial was retained for further analysis. The intensity measures were corrected by a calibration factor that was calculated using a sonometer and white noise, based on the guidelines of Maryn.[33] All audio recordings had a signal-to-noise ratio of at least 30 dB.

Multiparametric indices. The Dysphonia Severity Index (DSI) is a multiparametric approach designed to establish an objective and quantitative correlate of the perceived voice quality.[34] It is based on a weighted combination of the following parameters: MPT (s), F-high (Hz), I-low (dB), and jitter (%). The DSI is constructed as 0.13 MPT + 0.0053 F-high – 0.26 I-low – 1.18 jitter +12.4. The index ranges from −5 to +5 for severely dysphonic to normal voices. A more negative index indicates a worse voice quality. Values > +5 are possible in subjects with excellent vocal capacities. A DSI = +1.6 is the threshold separating normophonic from dysphonic persons. The DSI (v.02.02) script was used in Praat (v.6.2; [29]).

The Acoustic Voice Quality Index (AVQI) is an objective multiparametric approach to quantify dysphonia severity based on both a sustained vowel and continuous speech.[35] It consists of a weighted combination of six acoustic measures: three time domain measures (ie, shimmer local [SL], shimmer local dB [SLdB], and harmonics-to-noise ratio [HNR]), two frequency domain measures (ie, general slope of the spectrum [Slope] and tilt of the regression line through the spectrum [Tilt]), and one frequency domain measure (ie, smoothed cepstral peak prominence [CPPs]).[36] The formula of the index is 2.571 (3.295 – 0.111 CPPs – 0.073 HNR – 0.213 SL + 2.789 SLdB – 0.032 Slope + 0.077 Tilt) and ranges from 0 to 10. A higher index indicates a worse voice quality. The threshold score separating normophonic from dysphonic persons in Dutch is 2.95.[35] AVQI (v.02.03) was calculated on the audio recording of the sustained /a:/ vowel and the first two sentences of the text in Praat (v.6.2[29]).

Auditory-perceptual evaluation. The VAS of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V[37]) was used to evaluate the parameters grade, roughness, breathiness, asthenia, strain and instability.[38,39] The audio recordings of the /a:/ vowel and the text earlier used for the acoustic analysis were rated blindly by a speech-language pathologist (SLP) experienced in voice diagnostics (I.M.). Twenty percent of the samples were randomly repeated to determine intrarater reliability. The same 20% were also judged at random, blindly and independently, by another SLP experienced in voice diagnostics (C.L.) to measure interrater reliability.

Statistical analysis

SPSS version 27 (SPSS Corporation, Chicago, IL, USA) was used for the statistical analyses of the data. Analyses were conducted at $\alpha = 0.05$. Due to the high number of outcome measures, DSI, AVQI, f_o, G and VAS (vocal fatigue and complaints) were selected as primary outcomes for the analyses.

Linear mixed model analyses were performed to compare the evolution of the primary outcome measures between groups using the restricted maximum likelihood estimation and unstructured covariance structure. Time, Group, and Time x Group interactions were specified as fixed factors. A random intercept for subjects was included. Model assumptions were checked by inspecting whether residuals were normally distributed. Within-group effects of Time or Group were determined by posthoc pairwise comparisons.

Intraclass correlation coefficient (ICC) models were run to determine interrater (two-way mixed, consistency, average-measures) and intrarater reliability (a two-way mixed, absolute agreement, single-measures) of the auditory-perceptual data.

RESULTS

Voice-related information of the choir singers

Information regarding the participants' singing experience, vocal load, (history) of voice disorders and influencing factors can be found in Table 1.

TABLE 1.
Voice-related Information of the Choir Singers

	E (*n* = 15)	C (*n* = 15)	total (*n* = 30)
Singing experience and vocal load			
	n (%)	*n (%)*	*n (%)*
Amateur singer	14 (93.3)	14 (93.3)	28 (93.3)
Singing training	8 (53.3)	9 (60)	17 (56.7)
Professional voice user (occupation)	4 (28.6)	3 (20)	7 (24.1)
Vocal warm-up before singing	sometimes: 9 (60) always: 6 (40)	sometimes: 6 (40) always: 9 (60)	sometimes: 15 (50) always: 15 (50)
Vocal cool-down after singing	sometimes: 1 (6.7) always: 1 (6.7)	sometimes: 2 (13.3) always: 0 (0)	sometimes: 3 (10) always: 1 (6.7)
	M(SD; range)	*M(SD; range)*	*M(SD; range)*
Years of singing experience	19 (10.7; 6–50)	17 (7.4; 8–34)	18 (9.0; 6–50)
Hours of daily singing	1 (1.0; 0.5–4)	1 (0.6; 0.5–2.5)	1 (0.8; 0.5–4)
Hours of weekly singing	7 (7.0; 3–30)	7 (3.9; 3–15)	7 (5.6; 3–30)
(history) of voice disorders			
	n(%)	*n(%)*	*n(%)*
Voice disorder in the past	3 (20)	2 (13.3)	5 (16.7)
Current voice disorder	0 (0)	1 (6.7)	1 (3.3)
Phonosurgery in the past	0 (0)	1 (6.7)	1 (3.3)
Voice therapy in the past	2 (13.3)	2 (13.3)	4 (13.3)
Influencing factors			
	n (%)	*n (%)*	*n (%)*
Respiratory problems	2 (13.3)	2 (13.3)	4 (13.3)
Current cold/upper respiratory tract infection	3 (20)	2 (13.3)	5 (16.7)
Pregnancy	1 (6.7)	0 (0)	1 (3.3)
Smoking	1 (6.7)	2 (13.3)	3 (10)
Allergy	3 (20)	3 (20)	6 (20)
Reflux	4 (26.7)	6 (40)	10 (33.3)
Stress	13 (86.7)	10 (66.7)	23 (76.7)

Voice quality and PROMs of the choir singers

The baseline voice quality results and PROMs of the choir singers are presented in Table 2.

TABLE 2.
Voice Quality and PROMs of the Choir Singers (*n* = 30)

	M		*SD*		*Range*	
DSI	7.5		2.0		3.2–11.3	
AVQI	2.00		0.77		0.62–3.46	
f_o (Hz)	*men (n = 5)*	*women (n = 25)*	*men (n = 5)*	*women (n = 25)*	*men (n = 5)*	*women (n = 25)*
	99	213	16.8	25.4	69–110	173–268
G	5		7		0–24	
SVHI-10	14		4.5		4–25	
VTDS	26		12.6		0–52	
VFI	21		8.1		5–39	

Inter and intrarater reliability auditory-perceptual evaluation

Excellent interrater reliability was found for grade (ICC = 0.83), strain (ICC = 0.84) and instability (ICC = 0.81). Good interrater reliability was found for roughness (ICC = 0.60) and asthenia (ICC = 0.60). Fair interrater reliability was found for breathiness (ICC = 0.48). Concerning the intrarater reliability, excellent scores were found for grade (ICC = 0.80), roughness (ICC = 0.91), breathiness (ICC = 0.92) and instability (ICC = 0.75), and good scores were found for asthenia (ICC = 0.64) and strain (ICC = 0.60).[40]

Randomized controlled trial: effect of a performance on the choristers' voices

No significant differences were found between the two groups in age (Mann Whitney *U* test, $P = 0.493$), sex (Fisher's Exact test, $P > 0.999$) or any of the baseline outcome measures, indicating that randomization was successful. Results of the pre and post voice assessments for both groups can be found in Tables 3 (primary outcome measures) and Table 4 (secondary outcome measures).

TABLE 3.
Results of the Pre and Post Voice Assessments for the Experimental Group and the Control Group: Primary Outcome Measures

Outcome Measure	GROUP	TIME: Pre M	Pre CI	TIME: Post M	Post CI	LMM TIME* GROUP *P*-value
DSI	E	7.0	(6.0,8.1)	8.1	(7.0,9.3)	
	C	8.1	(7.0,9.2)	8.1	(6.8,9.3)	0.027*
AVQI	E	2.17	(1.77,2.57)	1.84	(1.41,2.26)	
	C	1.83	(1.41,2.24)	1.77	(1.33,2.20)	0.373
f_o	E	188	(162,215)	209	(179,239)	
	C	199	(173,226)	210	(181,241)	0.230
G	E	4	(0,7)	4	(1,6)	
	C	6	(3,10)	5	(2,8)	0.603
VAS	E	243	(148,338)	321	(231,410)	
	C	212	(113,310)	184	(91,277)	0.029*

* Indicates a significant effect ($P < 0.05$). *Abbreviations:* LMM, linear mixed model; E, experimental group; C, control group; M, mean; CI, confidence interval; DSI, Dysphonia Severity Index; AVQI, Acoustic Voice Quality Index; G, grade (auditory-perceptual evaluation); VAS, visual analogue scale.

TABLE 4.
Results of the Pre and Post Voice Assessments for the Experimental Group and the Control Group: Secondary Outcome Measures

Outcome measure	GROUP	TIME Pre M	Pre CI	Post M	Post CI
MPT	E	25.2	(21.0,29.4)	28.2	(23.7,32.8)
	C	25.2	(21.1,29.4)	25.6	(21.1,30.1)
Jitter	E	0.24	(0.18,0.30)	0.18	(0.13,0.24)
	C	0.20	(0.14,0.26)	0.20	(0.14,0.25)
F-high	E	963	(813,1113)	1010	(865,1156)
	C	1056	(907,1206)	1078	(932,1224)
F-low	E	121	(107,136)	125	(110,140)
	C	126	(112,141)	124	(109,139)
I-high	E	98	(96,101)	100	(97,102)
	C	99	(97,102)	100	(97,102)
I-low	E	52	(50,54)	51	(48,53)
	C	50	(48,52)	51	(49,54)
CPPS	E	14.54	(13.69,15.38)	14.86	(13.94,15.79)
	C	15.41	(14.54,16.29)	15.44	(14.48,16.39)
HNR	E	18.70	(17.01,20.38)	20.45	(19.05,21.84)
	C	20.02	(18.28,21.76)	20.88	(19.43,22.32)
Shimmer local	E	4.33	(3.66,5.01)	3.61	(2.93,4.28)
	C	3.66	(2.96,4.36)	3.62	(2.92,4.32)
Shimmer local dB	E	0.42	(0.38,0.47)	0.37	(0.33,0.42)
	C	0.38	(0.33,0.43)	0.39	(0.34,0.44)
Slope of LTAS	E	-16.28	(-17.96,-14.60)	-15.53	(-17.21,-13.85)
	C	-16.56	(-18.30,-14.82)	-16.12	(-17.85,-14.38)
Tilt of trendline through LTAS	E	-12.13	(-12.71,-11.55)	-11.94	(-12.52,-11.37)
	C	-12.10	(-12.70,-11.55)	-11.80	(-12.40,-11.20)
R	E	3	(1,6)	3	(0,7)
	C	4	(2,7)	5	(2,9)
B	E	1	(0,2)	1	(0,2)
	C	1	(0,3)	1	(0,2)
A	E	3	(0,6)	3	(1,6)
	C	4	(1,8)	3	(1,5)
S	E	3	(0,7)	3	(0,7)
	C	3	(0,6)	3	(0,7)
I	E	1	(0,3)	2	(0,3)
	C	2	(0,4)	1	(0,2)

Abbreviations: E, experimental group; C, control group; M, mean; CI, confidence interval; MPT, maximum phonation time; F-high, highest frequency; F-low, lowest frequency; I-high, highest intensity; I-low, lowest intensity; CPPS, smoothed cepstral peak prominence; HNR, harmonics-to-noise ratio; LTAS, long-term average spectrum; R, roughness; B, breathiness; A, asthenia: S, strain; I, instability.

Linear mixed model analysis showed a significant Time x Group interaction for the DSI ($P = 0.027$). Within-group posthoc tests revealed a significantly increased DSI for the experimental group (mean difference [MD] = +1.1, $P = 0.002$) and no change for the control group (MD = +0, $P = 0.962$). The evolution of the DSI can be seen in Figure 1. Within-group posthoc tests of the parameters of the DSI showed that the increment after singing was mainly due to a significant rise in MPT (MD = +3.0, $P = 0.001$) and F-high (MD = +47, $P = 0.036$), and a significant decrease in jitter (MD = -0.06, $P = 0.042$).

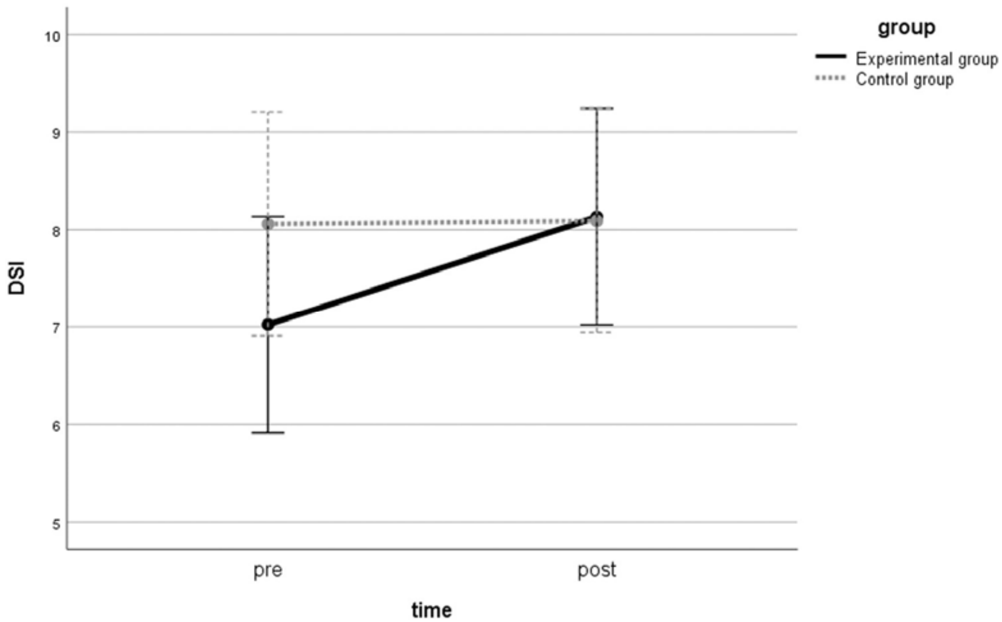

FIGURE 1. Evolution of the DSI in the experimental group and the control group.

There was also a significant Time x Group interaction for the VAS total score ($P = 0.029$) with a significant rise after singing in the experimental group (MD = +78, $P = 0.021$) and a non-significant decrease after the conversation in the control group (MD = -28, $P = 0.408$). The evolution of the VAS total score can be seen in Figure 2.

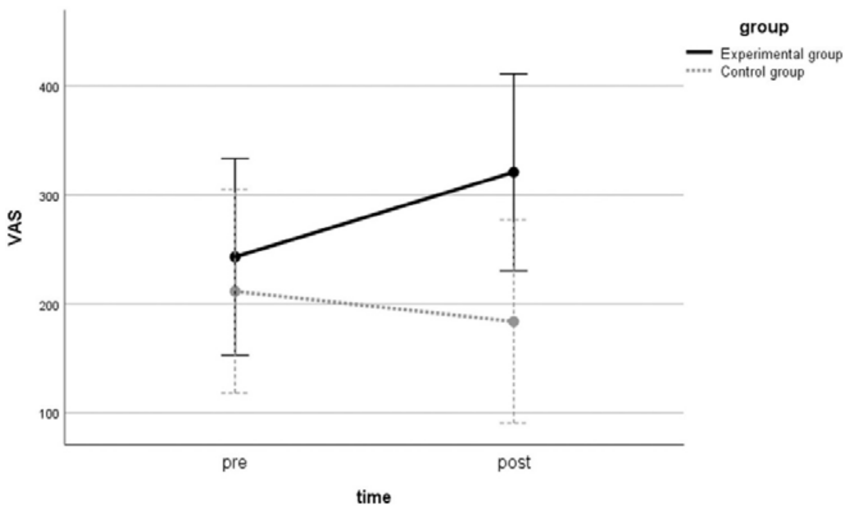

FIGURE 2. Evolution of the VAS in the experimental group and the control group.

No significant Time*Group interactions were found for the other primary outcome measures: AVQI, f_o or G. Within-group analyses showed a significant rise in f_o for both the experimental group (MD = + 21, $P = 0.001$) and the control group (MD = +11, $P = 0.044$), being more outspoken in the first.

DISCUSSION

The aims of this study were to determine the voice quality, capacities, symptoms and voice-related quality of life of choir singers, and to investigate the effect of a short performance, resembling warm-up duration (15 minutes), on the choristers' voices using a randomized controlled trial.

As expected, the choir singers showed excellent voice quality and capacities with mean DSI and AVQI scores of 7.5 (range 3.2–11.3) and 2.0 (range 0.62–3.46) respectively. The DSI scores are in line with the ones found in an earlier study examining choristers.[10] Each participant scored well above the threshold for DSI (> +1.6), and all but one scored below the threshold for AVQI (< 2.95). Auditory-perceptually, the mean grade score was 5/100 corresponding with a normal to mildly deviant voice quality.

Due to high vocal demands, often combined with limited vocal knowledge and training, choristers are at risk for voice impairments.[2, 3, 4, 5, 6, 7] This risk was visible in the PROMs with mean deviant scores on the SVHI-10 (14/40), VTDS (26/96) and VFI (21/96). PROMs are an indispensable part of a multidimensional voice assessment, especially for singers for whom small vocal deviations can have a tremendous impact.[7,41] The choral singers scored clearly above the mean SVHI value of 8.38 found in a large group of 528 vocally healthy singers,[25,41] indicating a considerable self-perceived singing voice handicap. Five participants (16.7%) actually reported a history of voice disorders, and one participant (3.3%) reported a current voice disorder. This number is comparable with the prevalence of 21% found by Ravall and Simberg.[2] However, in that study, a voice disorder was defined as experiencing two or more frequently occurring vocal symptoms, whereas no specific definition was used in the current study.

The choir singers seem vulnerable for stress with an occurrence rate of 76.7% (23/30). This number is comparable with the ones found in musical theater students.[42,43] Such high percentage is somewhat unexpected as choir singing is a leisure activity that has shown to reduce stress and positively affect the emotional state of the singers.[5,44] However, an important note here is that the competitive setting of the current study, ie the World Choir Games, can be seen as a stress-inducing situation and might have contributed to this high number. When interpreting the results, it should be taken into account that the high stress levels might have negatively affected voice quality.[45] Other common reported influencing factors were reflux (33.3%) and allergies (20%). Reflux was more frequently reported than in the studies of Ravall & Simberg[2] and Sharma et al.[6] who found percentages of only 8.3% and 15.8%, respectively. They did find identical percentages for allergies, upper respiratory tract infections and smoking.[2,6]

Choral singers can be seen as vocal athletes who need to prepare their body and mind for the specific vocal challenges of choir performance.[12,15] Nevertheless, to date, literature shows that vocal warm-up seems no embedded routine in choristers.[6,21] In the study of Sharma et al.[6], no less than 60% of the choir singers performed no vocal warm-up before singing. However, in the current study, vocal warm-up is clearly more represented, with half of the participants warming-up the voice prior to every singing performance and half of them warming-up sometimes. The most frequently reported warm-ups were scales (96.7%, 29/30) and pitch glides (86.7%, 26/30), followed by SOVT exercises (76.7%, 23/30), particularly humming, lip trills and tube phonation in water. Respiratory (66.7%, 20/30) and sigh exercises (13/30, 43.3%) are also frequently done by the participants.

Consistent with earlier studies,[13,15] a significant rise in f_o (+21Hz) was found after choir singing in the experimental group. A smaller (+11Hz), although still significant, rise was also found after 15 minute of talking in the control group. An increase in f_o can be seen as a condition of muscular warm-up, a normal physiological adaptation of the voice to speech or singing.[13,46] On the other hand, it can also be a marker of vocal resistance, effort or fatigue.[15,47,48] Regardless of the reason, a rise in f_o has been found after several vocal warm-up and training programs.[14,49, 50, 51] It has been considered as a result of muscle stretching and increased vocal fold tension, which is controlled by activation of the thyroarytenoid and cricothyroid muscles.[15,48,52,53] More tension in the intrinsic laryngeal muscles does not necessarily indicate less economic or efficient phonation, on the condition that the vocalist does not strain the laryngeal framework – such as cartilages, joints, and ligaments – to maintain adequate vocal fold tension and length.[51] In the end, however, it is important that the vocal musculature returns to its initial state after heavy vocal loading. Vocal cool-down can facilitate the process of muscle recovery and tissue regeneration for successive performances.[15] Nevertheless, vocal cool-down was reported by only four (13.3%) of the participants in the current study. A comparable low percentage (12.16%) was found in a large group (n = 526) of amateur choristers by Rosa and Behlau.[5] It is generally known that singers are less familiar with the value of vocal cool-down compared to warm-up.[13,18] This is an important point of attention to include in the education of (choir) singers.

The persistent uncertainty and discussion about whether a heavy vocal load, such as choir singing, induces warm-up or fatigue is reflected in the other results of the current study. Compared with the control group, the DSI significantly improved whereas the VAS (evaluating the presence of vocal fatigue and complaints) significantly deteriorated after singing. These findings might suggest that vocal efficiency and vocal fatigue can co-exist. The rise in DSI was due to an increase in MPT and F-high, indicating better vocal capacities and endurance after singing. Besides, there was a decrease in jitter indicating less acoustic perturbation and an improved voice quality. An increase in F-high and a decrease in jitter were also found after 60 minutes of choir singing in sopranos by Onofre et al.[13,15] Similarly, they report that self-perceived symptoms, such as hoarseness and fatigue, increased after singing for one hour.[13] Based on the limited literature available, it seems that 15 minute of choir singing induces comparable effects as 60 minutes of choir singing.[13,15] Longer choir singing, however, might be detrimental to the voice quality as Kitch et al.[22] reported a deterioration in terms of voice range and perturbation measures in tenors after 90 minutes. It should be noted that 20% of the participants in the current study show some vocal weakness (history of/current voice disorder) which might impact the group results on vocal fatigue and complaints after singing. It is possible that these self-perceived symptoms would not yet occur after 15minute of choir singing if stricter exclusion criteria were used. Nonetheless, the found percentage of vocal weakness seems representative for the general choir population,[2] and therefore probably provides an accurate indication of the real-life situation.

An important question now rises about whether or not a vocal warm-up program (apart from actual singing) before the choir performance would positively affect the PROMs and thus counteract the rise in self-perceived symptoms. To date, there is a lack of research investigating this matter. Manternach and Daugherty[54] investigated the effect of an SOVT warm-up with straw phonation on acoustic measures of an SATB (soprano, alto, tenor, and bass) choir and found that the overall spectral energy of the ensemble was increased after warm-up, indicating increased vocal output. This might be in line with the goal of SOVT exercises, i.e. obtaining vocal economy, or in other words, more output with less effort.[55, 56, 57] However, the phonatory effort was not investigated in this study and therefore no clear conclusions can be made. A later study of the same authors,[58] on the other hand, did include self-report measures and showed that most participants perceived the choir sound as better and reported a more efficient and comfortable voice production. Future studies should investigate the effect of different vocal warm-ups, both in terms of content and duration, on the voice quality and PROMs of choir singers.

This is a unique study, performed at the world largest international choir competition, that provides a multidimensional view of the voice of an understudied population of vocal athletes. Both PROMs, instrumental and auditory-perceptual outcomes were included. Furthermore, for the first time, the effect of a choir performance was investigated using a randomized controlled trial, which is considered the most reliable design for scientific evidence in efficacy studies. Randomization guarantees a similar distribution of risk factors in the groups of an experimental trial and reduce selection bias (Roland & Torgerson, 1998; Porzsolt et al., 2015). In randomized controlled trials, differences in outcome between experimental and control groups can be attributed to the tested intervention (Porzsolt et al., 2015).

Limitations of the current study are that the voice assessments could not be performed in a sound-treated room. However, a calibration factor was implemented, background noise was kept to a minimum and signal-to-noise ratios were all above 30dB. A second limitation is that there was no homogeneity of the singers in terms of voice type and singing style. Also, stricter protocols and standardization for both the experimental (e.g. standardized repertoire, strict instructions) and control condition (e.g. standardized questions) would have been a merit. Moreover, in future, time of data collection should be controlled relative to warm-ups, performances and vocal hygiene measures. At last, due to practical reasons, smaller subgroups of the choir performed instead of the entire choir, which might have led to an adaptation of their tessitura, blending and/or repertoire.

CONCLUSIONS

In conclusion, choir singers show excellent voice quality and capacities, that further improve after a short choir performance of 15 minutes. Vocal fatigue and complaints, on the other hand, already increase after that time span. Realizing that vocal load is much higher in real-life rehearsals, competitions and performances, choristers deserve and need a qualitative voice training and a strict follow-up. Future research can focus on effective vocal warm-up and cool-down for this population, and further explore the balance between vocal efficiency versus vocal fatigue.

Acknowledgments

This work was supported by a junior postdoctoral fellowship (I.M. / 1249721N) of the Research Foundation - Flanders (FWO).

REFERENCES

1. Ravi SK, Shabnam S, George KS, et al. Acoustic and aerodynamic characteristics of choral singers. J Voice. 2019; 33(5): 803.e1-803.e5.
2. Ravall S, Simberg S. Voice disorders and voice knowledge in choir singers. J Voice. 2020; 34:157.e1–157.e8.

3. Braun-Janzen C, Zeine L. Singers' interest and knowledge levels of vocal function and dysfunction: survey findings. J Voice. 2009;23:470–483.
4. Reddy BMS, Pebbili GK, Soonan VV. Smoothened cepstral peak prominence in choral singers, trained singers and non-singers. J All India Inst Speech Hear. 2014; 33:18–23.
5. Rosa M, Behlau M. Mapping of vocal risk in amateur choir. J Voice. 2017; 31: 118.e1–118.e11.
6. Sharma V, Srikanth N, Devadas U. A survey of vocal health in church choir singers. Eur Arch Otorhinolaryngol. 2021; 278:2907–2917.
7. Robotti C, Schindler A, Lechien JR, et al. Prevalence of laryngopharyngeal reflux symptoms, dysphonia, and vocal tract discomfort in amateur choir singers. J Voice. 2021. https://doi.org/10.1016/j.jvoice.2021.06.024.
8. Siupsinskiene N, Lycke H. Effects of vocal training on singing and speaking voice characteristics in vocally healthy adults and children based on choral and nonchoral data. J Voice. 2011; 25: e177–e189.
9. Irzaldy A, Wiyasihati SI, Purwanto B. Lung vital capacity of choir singers and nonsingers: a comparative study. J Voice. 2016; 30:717–720.
10. Berghs G, Creylman N, Avaux M, et al. A lifetime of professional singing: voice parameters and age in the Netherlands Radio Choir. Logoped Phoniatr Vocology. 2013; 38:59–63.
11. Miller R. Solutions for Singers. New York, NY: Oxford University Press; 2004.
12. Cook-Cunningham SL, Grady ML. The effects of three physical and vocal warm-up procedures on acoustic and perceptual measures of choral sound. J Voice. 2018; 32:192–199.
13. Onofre F, Ricz H, de Almeido Prado Y, et al. Vocal resistance among choir singers. Eur Arch Otorhinolaryngol. 2021; 278:159–165.
14. Van Lierde KM, D'haeseleer E, Baudonck N, et al. The impact of vocal warm-up exercises on the objective voice quality in female students training to be speech language pathologists. J Voice. 2011; 25:115–121.
15. Onofre F, Prado YA, Rojas GVE, et al. Measurements of the acoustic speaking voice after vocal warm-up and cooldown in choir singers. J Voice. 2017; 31: 129.e9–129.e14.
16. Elliott N, Sundberg J, Gramming P. What happens during vocal warm-up? J Voice. 1993; 9:37–44.
17. Sandage MJ, Wang S, Cao G. Upper airway thermoregulation during singing warm-up. J Voice. 2021; 35: 157.e1–157.e6.
18. Gish A, Kunduk M, Sims L, et al. Vocal warm-up practices and perceptions in vocalists: a pilot survey. J Voice. 2012; 26: e1–e10.
19. Ragsdale FW, Marchman JO, Bretl MM, et al. Quantifying subjective and objective measures of singing after different warm-up durations. J Voice. 2020. https://doi.org/10.1016/j.jvoice.2020.08.005.
20. Saldías M, Guzman M, Sandoval G, et al. Water resistance therapy as vocal warm-up method in contemporary commercial music singers. Folia Phoniatr Logop. 2020; 72:1–12.
21. Zoramsiami. Vocal Hygiene Practices in mizo Choral Singers [unpublished masters dissertation]. India: University of Mysore; 2014.
22. Kitch JA, Oates J, Greenwood K. Performance effects on the voices of 10 choral tenors: acoustic and perceptual findings. J Voice. 1996; 10:217–227.
23. Hunter EJ, Cantor-Cutiva LC, van Leer E, et al. Toward a consensus escription of vocal effort, vocal load, vocal loading, and vocal fatigue. J Speech Lang Hear Res. 2020; 63:509–532.
24. Cohen SM, Jacobson BH, Garrett CG, et al. Creation and validation of the Singing Voice Handicap Index. Annal Otol Rhinol Laryngol. 2007; 116:402–406.

25. Cohen SM, Statham M, Rosen CA, et al. Development and validation of the Singing Voice Handicap-10. Laryngoscope. 2009; 119:1864–1869.
26. Luyten A, Bruneel L, Meerschman I, et al. Prevalence of vocal tract discomfort in the Flemish population without self-perceived voice disorders. J Voice. 2016; 30:308–314.
27. Mathieson L, Hirani SP, Epstein R, et al. Laryngeal manual therapy: a preliminary study to examine its treatment effects in the management of muscle tension dysphonia. J Voice. 2009; 23:353–366.
28. Nanjundeswaran C, Jacobson BH, Gartner-Schmidt J, et al. Vocal Fatigue Index (VFI): development and validation. J Voice. 2015; 29:433–440.
29. Boersma W, Weenink D. Praat: Doing phonetics by computer (Version 6.2). 2021. Available at: http://www.praat.org/. Accessed October 31, 2021.
30. Van de Weijer J, Slis I. Nasaliteitsmeting met de nasometer [Nasalance measurement using the nasometer]. Logopedie en Foniatrie. 1991; 63:97–101.
31. Kohler K. German. J Int Phonetic Assoc. 1990; 20:48–50.
32. Fairbanks G. Voice and Articulation Drillbook. 2nd ed. New York: Harper & Row; 1960:124–139.
33. Maryn Y. The Acoustic Voice Quality Index in the programme Praat: a practical guideline. Belsele: Vlaamse Vereniging voor Logopedisten; 2013.
34. Wuyts FL, De Bodt MS, Molenberghs G, et al. The Dysphonia Severity Index: an objective measure of voice quality based on a multi-parameter approach. J Speech Lang Hear Res. 2000; 43:796–809.
35. Maryn Y, Corthals P, Van Cauwenberghe P, et al. Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. J Voice. 2010; 24:540–555.
36. Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. J Commun Disord. 2010; 43:161–174.
37. Kempster GB, Gerratt BR, Verdolini Abbott K, et al. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. Am J Speech Lang Pathol. 2009; 18:124–132.
38. Hirano M. Clinical Examination of Voice. New York, NY: Springer Verlag; 1981.
39. Dejonckere PH, Remacle M, Fresnel-Elbaz M, et al. Differential perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. Revue De Laryngologie - Otologie −Rhinologie. 1996; 117:219–224.
40. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess. 1994; 6:284–290.
41. Sobol M, Sielska-Badurek EM, Osuch-Wójcikiewicz E. Normative values for singing voice handicap index − systematic review and meta-analysis. Br J Otorhinolaryngol. 2020; 86:497–501.
42. D’haeseleer E, Claeys S, Meerschman I, et al. Vocal characteristics and laryngoscopic findings in future musical theater performers. J Voice. 2017; 31:462–469.
43. D’haeseleer E, Quintyn F, Kissel I, et al. Voice quality, symptoms, and habits in musical theater actors. J Voice. 2022; 36: 392.e1–392.e9.
44. Kreutz G, Bongard S, Rohrmann S, et al. Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state. J Behav Med. 2004; 27:623–635.
45. Van Lierde K, Van Heule S, De Ley S, et al. Effect of psychological stress on female voice quality. A multiparameter approach. Folia Phoniatr Logop. 2009; 61:105–111.
46. Rantala L, Vilkman E, Bloigu R. Voice changes during work: subjective complaints and objective measurements for female primary and secondary schoolteachers. J Voice. 2002; 16:344–355.

47. Vogel AP, Fletcher J, Maruf P. Acoustic analysis of the effects of sustained wakefulness on speech. J Acoust Soc Am. 2010; 128:3747–3756.
48. Pellicani AD, Ricz HM, Ricz LN. Phonatory function after prolonged voice use in Brazilian women. Codas. 2015; 27:392–399.
49. Meerschman I, D'haeseleer E, De Cock E, et al. Effectiveness of chewing technique on the phonation of female speech-language pathology students: a pilot study. J Voice. 2016; 30:574–578.
50. Meerschman I, Bettens K, Dejagere S, et al. Effect of two isolated vocal facilitating techniques chant talk and pitch inflections on the phonation of female speech-language pathology students: a pilot study. J Voice. 2016; 30: 771.e17–771.e25.
51. Meerschman I, D'haeseleer E, Catry T, et al. Effect of two isolated vocal facilitating techniques glottal fry and yawn-sigh on the phonation of female speech-language pathology studentes: a pilot study. J Commun Disord. 2017; 66:40–50.
52. Titze IR. Principles of Voice Production. Iowa City, USA: National Center for Voice and Speech; 2000.
53. De Bodt M, Heylen L, Mertens F, et al. Stemstoornissen: handleiding voor de klinische praktijk. Antwerpen: Garant; 2008.
54. Manternach JN, Clark C, Daugherty JF. Effects of a straw phonation protocol on acoustic measures of an SATB chorus singing two contrasting renaissance works. J Voice. 2017; 31: 514.e5–514.e10.
55. Titze IR, Laukkanen AM. Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. Logoped Phoniatr Vocol. 2007; 32:147–156.
56. Titze IR, Verdolini Abbott K. Vocology: The Science and Practice of Voice Habilitation. Salt Lake City, UT: National Center for Voice and Speech; 2012.
57. Guzman M, Laukkanen AM, Krupa P, et al. Vocal tract and glottal function during and after vocal exercising with resonance tube and straw. J Voice. 2013; 27: 523.e19–523.e34.
58. Manternach JN, Daugherty JF. Effects of a straw phonation protocol on acoustic and perceptual measures of an SATB chorus. J Voice. 2019; 33:86–88.