Spring 2023

The Effect of Different Warm-up Durations on Subjective and Objective Measures of Singing in Choral Singers Over the Age of 55

Jeremy A. Chesman
Missouri State University, JChesman@MissouriState.edu

THE EFFECT OF DIFFERENT WARM-UP DURATIONS ON SUBJECTIVE AND OBJECTIVE MEASURES OF SINGING IN CHORAL SINGERS OVER THE AGE OF 55

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Communication Sciences and Disorders

By

Jeremy A. Chesman

May 2023

THE EFFECT OF DIFFERENT WARM-UP DURATIONS ON SUBJECTIVE AND OBJECTIVE MEASURES OF SINGING IN CHORAL SINGERS OVER THE AGE OF 55

Communication Sciences and Disorders

Missouri State University, May 2023

Master of Science

Jeremy A. Chesman

ABSTRACT

Choral singing is a popular activity in the United States. Choral singers are often encouraged to warm up vocally before they sing. Considering voice conditions, like presbyphonia, that can develop shortly after retirement, more research about vocal warm-ups is needed for those over the age of 55. This study assesses the effects of various durations of vocal warm-ups on subjective and objective measures of the singing voice using a within-groups design with randomized condition order. Nine participants performed vocal warm-ups for 0, 5, 10, and 15 minutes. A song sample was then recorded and sent to two expert raters who rated the samples according to the Auditory-Perceptual Rating Instrument for Operatic Singing Voice. Participants also filled out the Evaluation of the Ability to Sing Easily (EASE) scale to rate how they felt about their singing. In addition to these two subjective measures, objective data were taken on pitch (Hz) and loudness (dB) of the highest loudest, highest softest, lowest loudest, and lowest softest pitches. There were no statistically significant changes in subjective or objective measures amongst the warm-up durations. This contrasts with a similar study performed on college-age music majors. Further research is needed to identify vocal warm-up types and durations that are effective with older adults.

KEYWORDS: singing, vocal pedagogy, vocal warm-up, voice, voice perception, choral singing

THE EFFECT OF DIFFERENT WARM-UP DURATIONS ON SUBJECTIVE AND OBJECTIVE MEASURES OF SINGING IN CHORAL SINGERS OVER THE AGE OF 55

By

Jeremy A. Chesman

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Communication Sciences and Disorders

May 2023

Approved:

Alana Mantie-Kozlowski, Ph.D., Thesis Committee Chair

Dee Telting, Ph.D., Committee Member

Sarah Lockenvitz, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to thank my thesis committee, Dr. Dee Telting, Dr. Alana Mantie-Kozlowski, and Dr. Sarah Lockenvitz for serving on my thesis committee. In addition, I would like to thank my colleagues Dr. Erin Plisco, Dr. Ann Marie Daehn, Dr. Paula Patterson, and Tyler Heckman for their assistance with subject recruitment and listening to and evaluating the recordings for this project.

Prof. Jennifer Pratt and Dr. Wafaa Kaf of the Communication Sciences and Disorders Department have served as informal mentors and sounding boards throughout my graduate program, and I thank them for the time they spent with me. Finally, my husband, Dr. Roberto Canales, has been an invaluable rock throughout the course of my studies. The literal least I could do for him is to acknowledge him in my thesis for all his support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Choral singing continues to be the most popular form of participation in the performing arts in America. While exact numbers of adult choral singers are difficult to obtain, there are an estimated 43 million adult choral singers in the United States, with at least 10 million estimated to be over the age of 65 (Grunwald Associates LLC & Chorus America, 2019). Maintaining vocal health is important for singers of all ages so they may continue to sing and participate fully in life activities.

Choral singers are generally encouraged to vocalize for a period before engaging in rehearsing or performing. There is a consensus that vocal warm-ups are beneficial to singers, but the mechanism of their benefit is not well understood (Hoch & Sandage 2018). The exercise physiology literature has found physiological, psychological, and bioenergetic effects of warm-ups. In athletes, studies have shown that physiological benefits of warm-ups include faster propagation of neural impulses and faster enzyme catalysis in muscles with temperature changes above 0.05 °C (Pearce et al., 2012; Brooks et al., 2005). Psychologically, improved performance after warm-ups is attributed to the re-establishment of muscle activation patterns required for motor planning as well as increased arousal prior to performing a task (McArdle et al., 2010). Warm-ups, however, do not only produce benefits. Bioenergetically, muscles use energy and enzymes for metabolism in both warm-ups and performance (Brooks et al., 2005). Thus, while warm-ups improve performance, excessive warm-ups can deplete biological resources. The speech motor system is unique in that it is particularly resistant to fatigue (Kent, 2004). Therefore, the exercise physiology literature may have limited application when referring to vocal warm-ups. The research specific to vocal warm-ups continues to develop, but more study is needed in such areas as content, sequencing, and duration of warm-ups.

Ragsdale et al. (2022) studied the effect of various durations of warm-ups on college-aged music majors. They found singers perceived a benefit to vocal warm-ups, but that there were no significant differences in the ratings of song recordings evaluated by expert listeners of recordings made after varied warm-up durations. The design of the present study is based on the Ragsdale study and responds to the recommendation to perform a similar study on a different population. The present study focuses on older adult choral singers. As the definition of "older adult" is imprecise, an age of 55 was selected. Recent research suggests that presbyphonia, a voice disorder that affects older adults, may be related to changes in amount of daily voice use that can occur with retirement (Ziegler & Hapner, 2020). While the average retirement age in the United States is 65 for men and 63 for women, the age of 55 was selected for this study as there will certainly be some members of the population who retire below the average age (Center for Retirement Research, n.d.).

This study examines the effects of various durations of vocal warm-ups on subjective and objective measures of singing in non-professional singers over the age of 55. The content and sequencing of the warm-ups are held constant, while the duration of the warm-up functions as the independent variable. This study aims to identify the ideal warm-up time for a choral singer over the age of 55, while also contributing to the overall literature on vocal warm-ups.

LITERATURE REVIEW

Objective Voice Measures

The objective effects of vocal warm-ups have been studied in various ways. Amir et al. (2004) studied 20 female singers with an average age of 18.62 years. In their study, singers recorded sustained /a/ and /i/ vowels at pitches 20%, 50%, and 80% of their reported vocal range. After the initial recording, singers warmed up using their own routine and duration. The mean duration of the singers in the study was 11 minutes (range: 7 to 23 minutes). They found that warm-ups reduced frequency and amplitude perturbation values, increased singer's formant amplitude, and improved noise-to-harmonic ratio (NHR). Amir et al. also found a difference in voice type in this study. The effect on perturbation measures was more significant with mezzo-soprano (lower female) voices and in lower pitch ranges.

McHenry et al. (2009) found that warm-ups affected measures of phonation threshold pressure (PTP), jitter, and NHR differently in males and females. That study included 10 males and 10 females ranging in age from 17 to 25 years. It examined the effect of a vocal warm-up with and without aerobic exercise. Jitter was reduced in males who performed warm-ups without an aerobic component. Females were found to have a greater reduction in PTP with the aerobic condition. Due to differences in objective measures between genders, research design in warm-ups should consider gender as one factor of analysis.

Subjective Measures

The study by McHenry et al. (2009) also included a subjective component in which participants rated their vocal effort. Both the males and the females demonstrated a reduction in perceived vocal effort, but the female group demonstrated a reduction of twice the magnitude of the male group. The authors hypothesized that the difference was due to differences in physical fitness levels, as evidenced by the average time each group took to attain their target heart rate with aerobic exercises (122 seconds for males and 88 seconds for females).

Self-perception of vocal warm-ups is a common factor studied in the literature. Gish et al. (2012) surveyed 117 professional and student singers on their vocal warm-up habits and perceptions. Of that total, 50 (42.7%) were age 41+, though no additional data on age was given. This study found that 5-note scales and glissandos were the most common singing exercises in a warm-up. Use of semi-occluded vocal tract exercises (SOVT) were reported by at least half of the participants, with lip/tongue trills preferred over nasal consonants. Singers agreed or strongly agreed almost unanimously that it is important to warm up before singing (95.7%). This study also found that singers felt their voice was more “cooperative” (92.3%) and that they felt more confident (95.7%) after vocal warm-ups. Based on survey data, the authors recommended a warm-up duration of 5-15 minutes.

Moorcroft and Kenny (2013) also used self-perception measures in their study, but they paired this with evaluations by experts of song recordings. They included 12 female singers with a mean age of 24.6 years. Singers were recorded before and after at 25-minute warm-up. Self-rating by the singers demonstrated unanimous improvement from pre- and post-intervention. The authors note this is particularly significant because warm-up duration and content were the same amongst participants, and this prescribed warm-up often differed from their reported preferred duration and content. This suggests that singers can perceive a benefit even with prescribed warm-ups, rather than using their traditional warm-up routine. That is, the psychological benefit of a warm-up is not only a function of the routine.

In addition to an improvement in self-perception, Moorcroft and Kenny also found an improvement in ratings by expert listeners. Recordings were assessed by singers themselves and by expert listeners. Thirteen categories were identified, such as mellowness, brilliance, vibrato, absence of strain. Both the performers and the listeners reported significantly improved scores after warm-ups. However, neither group consistently agreed on which of the individual categories improved. The instrument used was developed for this study (identification and definition of 13 singing qualities), and the study provided a reliability analysis. This analysis indicated that rate of vibrato before and after the warm-up affected the inter-rater reliability.

Instruments

There are a number of instruments that have been developed for self- and expert-assessment of singing voice. One tool for self-rating of singing is the Evaluation of the Ability to Sing Easily (EASE), developed by Phyland et al. (2013). This tool was developed for use with healthy voice users, utilizing ninety-five vocal health descriptors that were first identified from focus group interviews with 25 performing musical theatre (MT) singers. Using consensus techniques, these were reduced to 42 items. These items were then administered to 284 professional MT singers, of which 10 were in the age group 50-59 and none were above the age of 59. The results of that survey were analyzed for their fit with the Rasch model. The result is a 20-item scale that fits with the Rasch model (P = 0.04), no differential item functioning (DIF) by sex or age, and good internal consistency reliability.

The EASE scale was used in a follow-up study with amateur MT singers (Vella et al. 2021). This study used an updated version of the EASE scale with 22 items. The two additional items are part of the Vocal Concern (VC) subscale, with the questions “I am worried about my voice" and "I am concerned about my voice" (Appendix A). The 51 amateur singers in this study were in a musical theatre production at the time. They reported less vocal-use demands than in the first validation study with professional MT singers (Phyland et al. 2013). There were no significant differences in perceived vocal functioning related to sex or age and voice-related characteristics. Furthermore, there were no significant differences between this amateur cohortH and the professional cohort from the Phyland study. The Vella study supports the overall validity of the EASE scale and its use with amateur MT singers. Vella et al. recommend that the EASE scale continue to be investigated for broader clinical applications and with other populations.

Marchand et al. (2019) used the Brazilian version of the EASE scale (EASE-BR) with an amateur choir. In a prospective cross-sectional study, 44 singers completed the EASE scale before and after their regular vocal warm-ups of five minutes duration. A month later, participants completed the EASE scale before and after a warm-up "intervention," which consisted of 30 minutes of more extended warm-ups. The reported Cronbach alpha of the Total EASE score was 0.967, while the alphas of the voice concern (VC), Voice Fatigue (VF), and Voice Risk Index (VRI) were 0.914, 0.914, and 0.948, respectively. The authors concluded that the EASE scale was effective for use with amaetur choral singers.

In addition to self-rating scales, there are also validated instruments for listener perceptions. The Auditory-Perceptual Rating Instrument for the Operatic Singing Voice (APRIOSV) was developed as an objective tool for multidimensional evaluation of the operatic singing voice (Oates et al. 2006). Similar to the expert ratings in the Moorcroft and Kenny study (2004) above, the APRIOSV asks raters to give an overall score as well as scores on individual characteristics of the voice, such as ring, appropriate vibrato, and strain (Appendix B). Unlike the Moorcroft and Kenny study, the items on the APRIOSV have been psychometrically tested. Both an Equal-Appearing Interval (EAI) and a Visual Analog (VA) scale were examined in the study. Both versions of the scale were used in the same ways by evaluators, and evaluators did not express a clear perference for one scale over the other. The EAI scale appears to be more reliable, but the authors concluded that result is tentative because their study did not evaluate the relative sensitivity of the two forms. The population for the study was 21 professional opera chorus artists (13 female, 8 male). The size of the study is small to generate robust psychometric conclusions. However, the face validity is very good, and the available data support the use of the APRIOSV for listener evaluations. In the absence of a better instrument or a self-generated instrument, the APRIOSV provides a stastistically justifiable instrument for listener evaluation of classical singing.

In a study of 5 professional sopranos, the APRIOSV was used by external evaluators to describe the quality of two different high notes (A4, ≈ 880 Hz and D5, ≈1175 Hz). The evaluators were three speech therapists and three singing teachers. This information was paired with self-proprioceptive evaluations from the singers. While the study was limited due to its size and lack of clear research question, it demonstrates that the APRIOSV has been utilized in research involving listener ratings (Fernandes & Andrade E Silva 2020).

Statistics

While the instruments listed thus far have been psychometrically evaluated, careful statistical analysis is required when using them because they involve judges's rating. A variety of non-parametric statsistics have been developed to address the issues that come up with experiments that employ judges. One such test is the Friedman Test, sometimes called the Friedman Ranked Sums Test.

The Friedman test is a statistical test that is a non-parametric version of the a sum of squares treatment (SST). It uses a randomized block design that compares *b* blocks (participants) with *k* treatments. Within each block *b*, results of treatments are ranked from 1 (the smallest in the block) to *k* (the largest in the block). If two or more observations within each block are the same, the ranks that would have been assigned to these treatments are averaged. The ranks are then summed and squared, similar to an SST analysis. The statistic assumes that treatments are randomly assigned to experimental unites within blocks and that the number of blocks (*b*) or the number of treatments (*k*) exceeds 5.

Vocal Warm-Ups

Vocal warm-ups are a well-established pracitce for both choral and solo singing. There are many different types of vocal warm-ups, and warm-ups can vary in duration, content, and sequencing. There are a number of resources that contain different types of warm-ups, but much of the content is anecdotal. Empirical research on vocal warm-ups continues to develop with varaied results.

The exercise physiology literature indicates that an increase in muscle temperature from warm-ups helps improves biochemical processes of enzyme catalysis and propagate neural impulses faster (Pearce et al., 2012; Brooks et al., 2005). This has led researchers to investigate how physical warm-ups, either alone or with singing warm-ups, affects the voice. In a study of 10 males and 10 females from age 17 to 25, McHenry et al. (2009) examined the effect of a vocal warm-up with and without an aerobic component on both subjective and objectives measures. They found improvements in subjective measures to different degrees with both genders and improvements in different measures between the genders. This study is described in more detail above.

The use of physical warm-ups in choral singers has also been studied. In a pair of similar studies with college students comprising one study cohort, and elementary/high school students comprising the second cohort. Both studies used video-recorded warm ups with one of three conditions: physical only, vocal only, and physical/vocal combination. In the first study, there were a total of 61 participants comprised of both music majors and non-majors (Choir A: N=22, Mean Age = 22.68 years; Choir B: N=21, Mean Age=21.50 years; Choir C: N=18, Mean Age=21.00 years). The authors found that the warm-up that included physical and vocal aspects is preferred by singers, enables more resonant singing, and leads to more in-tune singing (Cook-Cunningham & Grady 2018). A second study by the same authors (2020) examined school-age choirs. In that study, there were a total of 140 participants (Choir D: N=45, Mean Age=10.40 years; Choir E: N=46, Mean Age=11.58 years; Choir F: N=49, Mean Age=14.81 years). This second study supported the results of the first study, indicating that physical and vocal warm-ups were preferred and led to improved intonation and greater resonance.

Another element that is often included in a vocal warm-up is a semi-occluded vocal tract (SOVT) exercise. In a study of 40 choristers between 18 and 57 years of age (average of 26.28 years), Cardoso et al. (2020) used a resonance tube to create an SOVT. One end of the tube was placed between the lips while the other end was placed in water. Participants were asked to perform *glissandi* (pitch glides) for three minutes ascending and three minutes descending. A Voice Range Profile (VRP) was taken before and after this exercise. A VRP is a clinical tool that creates a two-dimensional profile of the vocal range with the parameters of frequency and intensity (pitch and loudness). They found that the use of the resonance tube increased the VRP of alto, tenor, and bass voices, but did not find a significant result in soprano voices. They also found a change in maximum frequency, Hz range, and semitone range in all voice types after the use of the resonance tube.

Duke et al. (2015) also used SOVT exercises in their study of vocal warm-ups. They studied 13 male singers ranging in age from 19 to 42 (Mean Age=22.615 years) who sang in a variety of styles (classical, choral, jazz, gospel, opera). That study included three treatments: no warm-up, traditional warm-up (9 minutes duration), and warm-up with SOVT (6 minutes duration). The singing power ratio (SPR), which is the ratio of the highest peak between 2 kHz and 4 kHz and the highest peak between 0 kHz and 2 kHz, was calculated using four different vowels from a song recording. The perceived phonatory effort (PPE) was also determined by the singer with a visual analog scale. This study did not find any difference in SPR or PPE between any treatments.

Many studies have examined the effect of duration of vocal warm-up. While a 0-minute warm-up is often included in studies in order to have a baseline of function and to test the null hypothesis, some warm-up is recommended for improved performance and for injury prevention. Vinturri et al (2001) have reviewed the literature and found between 5 and 45 minutes of vocal warm-ups may be necessary to notice a difference in the voice. In a survey of 117 singers, the most common warm-up length was between 5 and 10 minutes. In that study, only one singer reported warming up longer than 30 minutes.

Ragsdale et al. (2022) indicated that 5 and 10 minutes of vocal warm-up increased the frequency range for females. In that study, 9 college-age music majors (6 females, 3 males) were recorded on four separate occasions. On each occasion, they warmed up for either 0, 5, 10, or 15 minutes. After warming up, they recorded a song sample and created a modified Voice Range

Profile (mVRP) that recorded their highest and lowerst frequencies at their loudest and softest intensities. The song sample was evaluated by expert raters using the Auditory-Perceptual Rating Instrument for the Operatic Singing Voice. Participants also completed the EASE scale after each session. The expert raters found no significant difference between warm-up durations, but did find an improvement in EASE scores at the 50 and 10-minute durations.

The current study extends the research by Ragsdale et al. (2022) by using siumilar methods to collect subjective and objective data on the effect of warm-up lengths on choral singers over the age of 55. It is hypothesized that the 5- and 10-minute warm-up durations will have no effect on perceptual analysis, but will have a positive effects on frequency tange and self-assessment of singing.

METHODS

This study utilized a within-groups design with randomized condition order. Participants scheduled five visits. The first visit consisted of an inclusion screening and providing informed consent. In the next four visits, participants warmed up for a duration of 0, 5, 10, or 15 minutes, which was pre-determined through randomization. They then sang a section of a song, which was recorded and later sent to expert raters for a subjective perceptual rating. Participants also took measures of the frequency and intensity (pitch and loudness) of their highest loudest comfortable sustainable pitch, as well as their highest softest, lowest loudest, and lowest softest comfortable sustainable pitches.

This study was approved by the Institutional Review Board on September 28, 2022 and received approval #IRB-FY2022-548 (Appendix C). Nine study participants were recruited from local choirs through flyers and emails from with the assistance of local choir directors. Qualifications for participation included speaking English, being over the age of 55, singing regularly (as defined by participation in musical rehearsal or perfroance 2 or more times per month).

Participants were screened for the presence of dysphonia using the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) and were excluded if their overall score was over 20. Additionally, a vocal health questionnaire was completed including information related to demographics, singing practice, voice-related medical history, and vocal hygiene habits. Participants were excluded if they had a history of diagnosed vocal pathology within the last year. Participants were also excluded if they were transgender, due to potential effects that hornomes have on the voice.

A script with standardized instructions was used when collecting data from participants. All data were collected by the primary investigator. Participants were engaged in study-related activities for approximately 4 weeks. Participants met with the primary investigator once or twice a week for 5-35 minutes. Days of the week and time of day for data collection were kept consistent. Participants were asked to not sing, warm-up their voice, or extensively talk the day of their data collection session.

Warm-up Time

During data collection sessions, participants warmed up their voices by following recorded voice warm-ups. This standardized warm-ups across participants. Warm-up exercises consisted of vowel and consonant combinations across scales and arpeggios, tongue twisters and syllable repetions for engagement of articulators, and semi-occluded vocal tract exercises. Each participant warmed up for either 0, 5, 10, or 15 minutes (Appendix D). Each warm-up duration was completed one time by each participant over the course of the study. The order of warm-up duration was randomized using Microsoft Excel. The range of time in between sessions for each participant was 3-10 days.

Song Recording

Immediately following the warm-up, participants were recorded singing “America the Beautiful” in an appropriate key. Participants selected the key and familiarized themselves with the song prior to the initial data collection session. They were instructed to be able to sing through the piece without musical or text errors. The key was consistent per participant for all recordings, and all participants selected either B-flat major or C major for the song recording. A piano accompaniment was played on headphones for the participant while their singing sample was recorded, but the headphones presented the accompaniment from being heard on the recording. A Zoom H2n Handy Recorder was used to record the song. Mouth-to-microphone distance was set at 15 cm, which was measured before each item recorded. The recording was made in a soundproof audiological testing booth in the Communication Sciences and Disorders Department at Missouri State University.

Immediate Objective and Subjective Measures

After the song recording, a modified voice range profile (mVRP) was completed. Participants were asked to sing a scale or a pitch slide up to the highest loudest note they could comfortably sustain. A pitch-analyzing application (ClearTune) was used on a smartphone (iPhone SE) to collect frequency data. A handheld digital sound level meter (Radio Shack Digital Sound Level Meter) was used to record intensity data. Both were held at a mouth-to-microphone distance of 15 cm. After the frequency and intensity of the highest loudest pitch was recorded, the same procedure was followed for the highest softest, then the lowest loudest and lowest softest pitch.

The Evaluation of the Ability to Sing Easily (EASE) scale was completed after the mVRP (Appendix A). Participants were instructed to respond to the questionnaire as if the questions were about their performance during the song recording. After completion, the total EASE score (0-66) was calculated.

Delayed Perceptual Analysis

Audio samples were edited to remove the first phrase of the piece, approximately the first 20 seconds. Two independent blinded expert raters evaluated the audio recordings of all participants using the Auditory-Perceptual Rating Instrument for the Operatic Singing Voice. The expert raters were both university-level singing voice professors teaching. Recordings were scored on the 8 categories of the APRIOSV on a scale of 0-10, according to the instructions from the instrument (Appendix B). The order of presentation of audio samples was randomized for each evaluator using Microsoft Excel. After randomization, the order was verified by hand to ensure that no evaluator listened to the same participant twice in a row. Therefore, the sample is pseudo-randomized.

Expert raters were asked to use the same headphones for all ratings and ensure they were in a quiet environment that was free from distractions. Experts were instructed to listen to the recordings in their entirety prior to making their rating. They were also asked to complete all ratings in one sitting. Experts were blinded to warm-up durations and all other information about participants.

RESULTS

Demographic Data

Nine participants were recruited and screened for study inclusion. All participants had CAPE-V scores below 20 and met other study criteria for age, gender, and history of vocal pathology. The study included five men and four women. Three participants reported distant smoking history. Five participants reported drinking caffeinated beverages, with 2 being the median and four being the maximum. All participants reported choral singing experience of several years. Four of the nine participants reported warming up regularly before singing. Demographic and vocal health history are listed in Table 1.

Table 1. Demographic Data and Vocal Health History

Gender	Voice Type	Age	Years Voice Lessons	Years Choral Singing	Weekly Hours Solo Practice	Weekly Hours Group Rehearsal	Duration of Warm-up	Daily Water Intake (Oz.)
Male-5	Sop.-2	69	4	30	1	1	5	48
Female-4	Alto-2	[63.5,73.5]	[1,5.5]	[23,54]	[0,3.5]	[1,2.25]	[1.5,13.75]	[14,60]
	Tenor-2							
	Bass-3							

Gender and voice type reported by count.
Other variables reported by median and interquartile range

Immediate Perceptual Outcome

The Friedman rank sum test for total EASE score for all participants was not significant (P = 0.825). The interquartile ranges overlap for all warm-up durations, confirming that there is no significant change in EASE score among the four warm-up conditions (Fig. 1). The Friedman rank sum test requires the number of blocks (participants) or the number of treatments (warm-up duration) to be greater than five in order to keep statistical power. Therefore, all Friedman tests will be reported in total, with both genders combined into one group.

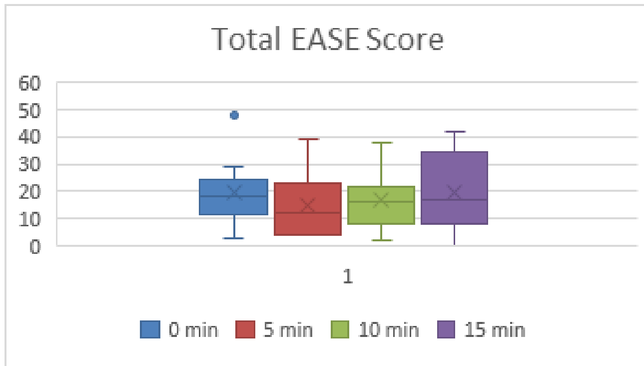

Figure 1. Boxplots of the medians and interquartile ranges for total EASE scores for all participants.

Delayed Perceptual Outcome

The average ratings for each category of the Auditory-Perceptual Rating Instrument for Operatic Singing Voice are included in Table 2. The table is organized by the aggregate mean of all the raters per duration condition. The final tow indicates the rnage (maximum value minus minimum value) for each of the categories of the scale. No category has a range greater than 1, suggesting that there was no significant difference in the perception of the song sample due to warm-up duration. Calculated using the median, the interquartile ranges also suggest no significant difference.

Table 2. Averages from scores of expert raters for the Auditory-Perceptual Rating Instrument for Operatic Singing Voice.

Length	Appropriate Vibrato	Resonance Balance	Ring	Pitch Accuracy	Breath Management	Evenness Throughout the Range	Strain	Overall Vocal Performance
0 min	5.556	5.778	5.833	7.167	6.611	6.500	6.111	6.167
5 min	5.111	5.611	5.444	7.333	6.389	6.167	6.111	5.833
10 min	5.611	5.722	5.500	7.667	6.500	6.778	5.944	6.167
15 min	5.889	5.944	5.611	7.500	6.833	6.500	6.556	6.389
Range	0.778	0.333	0.389	0.500	0.444	0.611	0.611	0.556

Modified Voice Range Profile

The results of the mVRP can be found in Figure 2. Friedman rank sum tests were performed on both the pitch (Hz) and intensity (dB) for the highest loudest, highest softest, lowest loudest, and lowest softest productions. Of these data point, the intensity of the highest loudest pitch was the only measure that was significant (P = 0.026). However, the interquartile ranges of the boxplot overlap, indicating no significant change amongst the four warm-up durations.

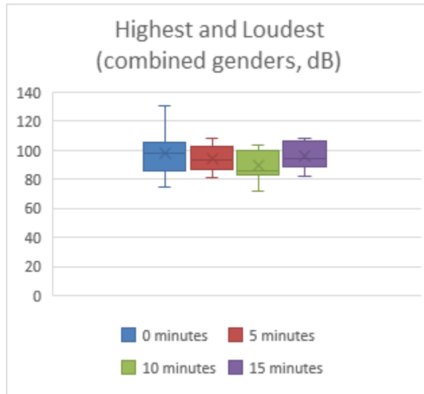

Figure 2. Boxplots of the medians and interquartile ranges for decibels of highest and loudest pitch.

DISCUSSION

The purpose of this study was to examine the effects of various warm-up durations on subjective and objectives measures of singing. On the subjective measures (the participants' self-rating on the EASE scale and the blinded experts' ratings on the Auditory-Perceptual Rating Instrument for the Operatic Singing Voice), there was no significant difference. On the objective measures, there was also no significant difference.

Previous to this study, Ragsdale et at. (2022) did a similar study with college-aged music majors. Both studies examined the objective factors of pitch and intensity at the extremes of the vocal range for four warm-up durations of 0, 5, 10, and 15 minutes. Both studies also incorporated subjective ratings. The current study did not included a self-assessment by the singers using the Auditory-Perceptual Rating Instrument for the Operatic Singing Voice that the Ragsdale study used, because the participants did not qualify as expert raters.

The Ragsdale study observed that female singers had a significant increase in range at the 5- and 10-minute warm-up durations, but observed no significant change in the range of male singers. However, the Ragsdale study had only three male participants over the four treatments. The Friedman test, however, requires at least five participants or five treatments. Given the number of participants, the Friedman test in the Ragsdale study did not have enough statistical power to find significance.

The current study also does not have sufficient statistical power to analyze the participants by gender, but does satisfy the conditions to look at the group as a whole. There were no significant differences in objective factors between any of the warm-up durations when all participants were analyzed in the aggregate. As the Ragsdale study does not report with genders combined, and the current study cannot meaningfully report with genders separated, no further conclusions can be made about objective measures.

There were no significant changes in the ratings of expert listeners due to warm-up duration. This is consistent with the conclusion from the Ragsdale study. The two studies differ in the results of the EASE scale scores. The current study found no significant change in the EASE score due to any warm-up duration ($P = 0.825$). The Ragsdale study found that the 5- and 10-minute warm-up durations did change the EASE score. The results in both studies were based on the complete sample of nine participants, and therefore meet the statistical assumptions of the Friedman test.

Limitations

There are several limitations to this study. Because of the small sample size, the Friedman Rank Sum Test was used. As discussed above, there is sufficient power to analyze the data in the aggregate, but not by gender. Furthermore, because the Friedman Test uses ranks of data, the magnitude of differences is ignored. Another statistical analysis may give more information about differences.

No hearing test was performed on participants or on evaluators. Auditory feedback could influence the perception of singers, so the lack of a hearing screening could have affected their perceptions. The same is true for evaluators who listened to the recordings. Additionally, each evaluator used the same set of headphones for perceptual analysis, but each evaluator used a different set of headphones from each other. This also may have affected perceptual analysis.

The Auditory-Perceptual Rating Instrument for the Operatic Singing Voice was normed on professional opera choral singers. Opera chorus singing is significantly different from other choral singing. Therefore the results from that instrument have not been validated for this population.

Future Directions

This study should be replicated with a larger number of participants. With six participants of each gender, the data analysis could be separated by gender. Currently, neither study can clearly comment on the effect of gender on warm-ups, particularly on the effect of warm-ups on males. A larger sample size may allow for the data to be analyzed with more powerful methods, as opposed to the ranking of data as in the Friedman Test.

Another area for future study would be the inclusion of an intervention protocol in the current study. Two singers reported weekly solo vocal practice of five hours per week, while the median was one hour. Of those two singers who practiced regularly, one did have a reduction in EASE score after vocal warm-up. The other had a reduction and an increase in EASE score on different warm-up durations. Examining the effects of vocal warm-up durations before and after the implementation of a regular vocal exercise protocol could give insight into the maintenance of vocal health and preferred occupations. This may have clinical correlations as well, given the research on intervention in presbyphonia (Ziegler and Hapner, 2020).

Conclusions

Singers over the age of 55 may not perceive a benefit to vocal warm-ups. While the college-aged students in the Ragsdale study did report a significant reduction in EASE scores, singers over the age of 55 did not. This has implications for choral rehearsal and performance. Since this population may not perceive a benefit to vocal warm-ups, they may avoid completing them before singing. Indeed, roughly half of the participants in the study reported not warming-up regularly. However, the general consensus in both the vocal warm-up literature as well as the exercise science literature is that warm-ups are beneficial for injury prevention. This study did not examine injury prevention, and therefore can draw no conclusions about the relationships between warm-ups and vocal health. However, from a performance standpoint, there appear to be no changes to range, volume, or perception of singing due to vocal warm-ups in singers of the age of 55.

REFERENCES

Amir, O., Amir, N., & Michaeli O. (2005). Evaluating the influence of warmup on singing voice quality using acoustic measures. *Journal of Voice, 19(2)*, 252-260.

Brooks, G.A., Fahey, T.D., & Baldwin, K.M. (2005). *Exercise Physiology: Human Bioenergetics and Its Applications* (4th ed.). McGraw-Hill.

Cardoso, N.S.V., Lucena, J.A., & Gomes, A.O.C. (2020). Immediate effect of a resonance tube on the vocal range profile of choristers. *Journal of Voice, 34(5),* 667-674.

Center for Retirement Research at Boston College. (n.d.) *Average Retirement Age for Men and Women, 1962-2016* [Infographic]. Retrieved Feb. 18, 2023, https://crr.bc.edu/wp-content/uploads/2021/03/Average-retirement-age_2017-CPS.pdf

Cook-Cunningham, S.L. & Grady, M.L. (2018). The effects of three physical and vocal warm-up procedures on acoustic and perceptual measures of choral sound. *Journal of Voice, 32(2),* 192-199.

Duke, E., Plexico, L.W., Sandage, M.J., & Hoch, M. (2015). The effect of traditional singing warm-up versus semi-occluded vocal tract exercises on the acoustic parameters of singing voice. *Journal of Voice, 29(6),* 727-732.

Fernandes, D..E., & Andrade E Silva, M.A. (2020). Upper high notes: auditory-perceptual analysis of voice and self-reporting among professional sopranos, *CoDAS, 32(4),* 1-6. https://doi.org/10.1590/2317-1782/20192019068

Gish, A., Kunduk, M. Sims, L., & McWhorter, A.J. (2012). Vocal warm-up practices and perceptions in vocalists: a pilto survey. *Journal of Voice, 26(1),* e1-e10.

Grady, M.L. & Cook-Cunningham, S.L. (2020). The effects of three physical and vocal warm-up procedures on acoustic and perceptual measures of choral sound: study replication with younger populations. *Journal of Voice, 34(4),* 547-647.

Grunwald Associates LLC & Chorus America. (2019). *The Chorus Impact Study: Singing for a Lifetime*.

Hoch, M. & Sandage, M.J. (2018). Exercise science principles and the vocal warm-up: Implications for singing voice pedagogy. *Journal of Voice, 32(1),* 79-84.

Kent, R.D. (2004). The uniqueness of speech of among motor systems. *Clinical Linguistics & Phonetics, 18(6-8),* 495-505.

Marchand, D.L.P., Kavaliunas, F.S., & Cassol, M. (2019). The effectiveness of the EASE scale in the development of a vocal warm-up program for an amateur choir. *Journal of Voice, 33(3),* 310-316.

McHenry, M., Johnson, J., & Foshea, B. (2009). The effect of specific versus combined warm-up strategies for the voice. *Journal of Voice, 23(5),* 572-576.

Moorcroft, L., & Kenny, D.T. (2013). Singer and listener perception of vocal warm-up. *Journal of Voice, 27(2),* 258.e1-258.e13.

Oates, J.M., Bain, B., Davis, P., Chapman, J., & Kenny, D. (2006). Development of an auditory-perceptual rating instrument for the operatic singing voice. *Journal of Voice, 20(1),* 71-81.

Pearce, A.J., Rowe, G.S., & Whyte, D.G. (2012). Neural conduction and excitability following a simple warm up. *Journal of Science & Medicine in Sport, 15(2),* 164-168.

Phyland, D.J., Pallant, J.F., Benninger, M.S., Thibeault, S.L., Greenwood, K.M., Smith, J.A., & Vallance, N. (2013). Development and preliminary validation of the EASE: A tool to measure perceived singing voice function. *Journal of Voice, 27(4),* 454-462.

Ragsdale, F.W., Marchman, J.O., Bretl, M.M., Diaz, J., Rosow, D.E., Anis, M., Zhang, H., Landera, M.A., & Lloyd, A.T. (2022). *Journal of Voice, 36(5),* 661-667.

Vella, B., Brown, L., & Phyland, D. (2021). Amateur music theatre singers' perceptions of their current singing voice function. *Journal of Voice, 35(4),* 589-596.

Vintturi, J., Alku, P., Lauri, E., Sala, E., Sihvo, M., & Vilkman, E. (2001). Objective analysis of vocal warm-up with special reference to ergonomic factors. *Journal of Voice, 15(1),* 36-53.

Ziegler, A. & Hapner, E. (2020). Vocal Dose in Older Adults with Presbyphonia: An Analytic, Cross-Sectional Study. *Journal of Voice, 34(2)*, 221-230. https://doi.org/10.1016/j.voice.2018.09.005

APPENDICES

Appendix A: Evaluation of Ability to Sing Easily (EASE) Scale

Scale: Not at all (0) Mildly (1) Moderately (2) Extremely (3)
*** Item is reverse scored

1. My voice is husky
2. My voice is dry/scratchy
3. My voice cracks and breaks
4. My throat muscles are feeling over-worked
5. My voice is breathy
6. My singing voice feels good***
7. The onsets of my notes are delayed or breathy
8. My voice feels strained
9. I am worried about my voice
10. I am having difficulty with my breath for long phrases
11. My top notes are breathy
12. My voice sounds rich and resonant***
13. My voice is cutting out on some notes
14. I am having difficulty singing softly
15. My voice is tired
16. I am having difficulty changing registers
17. I am having difficulty with my high notes
18. Singing feels like hard work
19. I am having difficulty projecting my voice
20. I am concerned about my voice
21. My voice feels ready for performance if required***
22. I am having difficulty sustaining long notes

Appendix B: Auditory-Perceptual Rating Instrument for the Operatic Singing Voice

Please circle a single number from 0-10 to indicate voice quality on each of the following:

Overall Vocal Performance (an overall rating of the aesthetic and technical quality of singing voice)

Poor 0 1 2 3 4 5 6 7 8 9 10 Excellent

Appropriate Vibrato (regular and smooth undulation of frequency of the tone)

Poor 0 1 2 3 4 5 6 7 8 9 10 Excellent

Resonance Balance (*chiaroscuro*) (appropriate balance of dark and light colors in the voice)

Poor 0 1 2 3 4 5 6 7 8 9 10 Excellent

Ring (brilliance of tone)

Poor 0 1 2 3 4 5 6 7 8 9 10 Excellent

Pitch Accuracy (singing in tune)

Poor 0 1 2 3 4 5 6 7 8 9 10 Excellent

Breath Management (efficient breath management)

Poor 0 1 2 3 4 5 6 7 8 9 10 Excellent

Evenness Throughout the Range (ability to sing freely throughout the pitch and dynamic range without inappropriate change in voice quality)

Poor 0 1 2 3 4 5 6 7 8 9 10 Excellent

Strain (voice quality that gives the impression of excessive vocal effort)

Severe Strain 0 1 2 3 4 5 6 7 8 9 10 Free from Strain

Appendix C: IRB Approval Letter

Missouri State
UNIVERSITY

To:
Jeremy Chesman
Music, Communication Sciences & Disor
Dee Telting

RE: Notice of IRB Approval
Submission Type: Initial
Study #: IRB-FY2022-548
Study Title: The Effect of Different Warm-Up Durations on Subjective and Objective Measures of Singing in Non-Professional Adults Over the Age of 55
Decision: Approved

Approval Date: September 28, 2022

This submission has been approved by the Missouri State University Institutional Review Board (IRB). You are required to obtain IRB approval for any changes to any aspect of this study before they can be implemented. Should any adverse event or unanticipated problem involving risks to subjects or others occur it must be reported immediately to the IRB.

This study was reviewed in accordance with federal regulations governing human subjects research, including those found at 45 CFR 46 (Common Rule), 45 CFR 164 (HIPAA), 21 CFR 50 & 56 (FDA), and 40 CFR 26 (EPA), where applicable.

Researchers Associated with this Project:
PI: Jeremy Chesman
Co-PI: Dee Telting
Primary Contact: Jeremy Chesman
Other Investigators: Ann Marie Daehn, Paula Patterson

Appendix D: Content of Pre-recorded Warm-ups

The recorded warm-ups for the 10-minute duration included all of the recordings for the 5-minute duration with additional exercises. The recordings for the 15-minute duration included all of the recordings from the previous two durations with additional exercises to total 15 minutes.

5 minutes	10 minutes	15 minutes
Descending pitch slides of a 5th on /u/	Descending pitch slides of a 5th on /u/	Descending pitch slides of a 5th on /u/
/bʌdʌgʌ/ descending scales in thirds	/uiuiuiuia/, descending 5-note pattern, ascending	/uiuiuiuia/, descending 5-note pattern, ascending
Ascending and descending siren on /u/	/bʌdʌgʌ/ descending scales in thirds	On /a/, ascending-descending 5-note pattern followed by ascending-descending octave scale, overall ascending
/mɑmɛmimomu/ up and down a fifth, ascending	Ascending and descending siren on /u/	/bʌdʌgʌ/ descending scales in thirds
/siŋ/ descending triad arpeggio	/bidi/, descending-ascending-descending five-note pattern, overall descending	Ascending and descending siren on /u/
Three descending and ascending arpeggios on /u/, /i/, /a/, descending	/mɑmɛmimomu/ up and down a fifth, ascending	/bidi/, descending-ascending-descending five-note pattern, overall descending
	"No no no no no," descending 5-note pattern, descending, in chest voice	Sliding up and down an octave on /a/, ascending
	/siŋ/ descending triad arpeggio, singing on /ŋ/	/mɑmɛmimomu/ up and down a fifth, ascending
	Lip trills- ascending octave leap followed by descending octave scale, overall ascending	"No no no no no," descending 5-note pattern, descending, in chest voice
	Three descending and ascending arpeggios on /u/, /i/, /a/, descending	Singing scale degrees 18531 on /a-ɛ-i-o-u/, ascending
	Scale degree 154535251 on /u/, ending on /a/, descending	/siŋ/ descending triad arpeggio, singing on /ŋ/
		Lip trills- ascending octave leap followed by descending octave scale, overall ascending

		Diadiokinetic syllable repetition and tongue twisters
		Three descending and ascending arpeggios on /u/, /i/, /a/, descending
		Scale degree 154535251 on /u/, ending on /a/, descending
		Staccato arpeggio up and down on /a/, ascending