Reducing Vocal Fatigue While Preserving Realism During Video Game Voice-Overs Using the Vocal Combat Technique: A Randomized Controlled Trial

[*]Katelyn Reid, [*,†,‡]Victoria S. McKenna, [§]C. Andrew Lee, [§]John Paul Giliberto, and [¶]D.'Arcy Smith, *†‡¶*Cincinnati, Ohio, and* §*Seattle, Washington*

Summary: Objective. Vocal Combat Technique (VCT) teaches indirect and direct behavioral voice techniques to voice-over artists performing in violent video games. Although previous work on VCT has shown promise for mitigating dysphonia symptoms, a randomized clinical trial has yet to be undertaken. Therefore, we completed a randomized, controlled trial between a group of experienced video game voice-over actors receiving VCT and a control group comparison.

Methods. A total of 24 video game voice-over actors completed this study. Participants were randomly assigned to receive VCT or indirect vocal hygiene training prior to completing an intensive 1-hour video game voice recording session. The primary outcome was a change in Voice Handicap Index-10 (VHI-10) preperformance/postperformance. Secondary measures included a modified version of the Evaluation of the Ability to Sing Easily (m-EASE), the Vocal Tract Discomfort Scale (VTDS), and questions regarding return to work. Participants were also rated on the realism of their vocal performance by a blinded video game director.

Results. The VCT group showed a significantly smaller change in VHI-10 and m-EASE scores postperformance, and a higher increased likelihood to return to work compared to the control group. There were no group differences for VTDS or realism ratings. Four participants from the control group exhibited outlier behavior with more pronounced phonotraumatic symptoms following performance than all other participants.

Conclusions. VCT shows evidence of mitigating symptoms of dysphonia while preserving the realism of the vocal performance. More work is needed to understand performers at risk for more severe vocal symptoms following extreme voice-over work, so as to target them for preventative techniques and voice preservation.

Key Words: Vocal combat technique–Vocal fatigue–Vocal performance–Video game voice-over.

INTRODUCTION

Video gaming is a $93-billion-per-year worldwide industry,[1] which continues to grow at an exponential rate. The video gaming industry is now exceeding the revenue of the film and music industries combined.[2] Within the gaming industry, "Shooter" and "Action Adventure" video games were the top two video game genres worldwide in 2022.[3] These video games require vocalizations to accompany character narratives and experiences. This has led to an increased demand for voice-over actors to perform extreme vocalizations (eg, shouting and screaming) and heavier workloads to make a living in the industry.[4]

It is well known that performers and professional voice users are susceptible to vocal pathology and dysphonia due to the duration and intensity of their occupational voice use.[5–7] Phonotrauma, in particular, has been found to be the leading cause of dysphonia for those involved in vocal performance.[8] Video game voice-over actors are one subgroup of vocal performers who are asked to engage in repetitive reproduction of overt phonotraumatic behaviors in order to emulate realism for characters experiencing exertion, danger, or pain.[9] Such behaviors include, but are not limited to, shouting, yelling, howling, and screaming; projecting voice as if over loud ambient noises, such as gunfire, explosions, and battle sounds; voice modifications of raspiness, harshness, roughness, hoarseness, and growl; and emulating violent events, such as suffocation, strangulation, and electrocution.[9]

A survey conducted in 2021 by the Alliance of Canadian Cinema, Television and Radio Artists (ACTRA) found that 74% of voice-over actors reported that their sessions included loud/projected, aggressive, or vocally extreme work "very often" or "almost always" and 38% of performers experienced vocal fatigue or stress during their sessions with the same frequency.[10] Most notably, 19% reported that they found it difficult to recover their normal vocal quality after a session, with 43% stating that it could take two or more days to return to baseline. More than one-quarter of performers (28%) revealed that they considered turning down a session for fear of the impact that it would have on their voice or the work that it would cause them to lose. The results of this survey indicate an immediate need for behavioral voice techniques that will

Accepted for publication July 6, 2023.
This work was supported by the University of Cincinnati College-Conservatory of Music Faculty Development Fund awarded to author D.S.
From the *Department of Otolaryngology—Head and Neck Surgery, University of Cincinnati, Cincinnati, Ohio; †Department of Communication Sciences and Disorders, University of Cincinnati, Cincinnati, Ohio; ‡Department of Biomedical Engineering, University of Cincinnati, Cincinnati, Ohio; §Department of Otolaryngology—Head and Neck Surgery, University of Washington, Seattle, Washington; and the ¶Department of Acting, College-Conservatory of Music, University of Cincinnati, Cincinnati, Ohio.
Address correspondence and reprint requests to Katelyn Reid, Department of Otolaryngology—Head and Neck Surgery, University of Cincinnati, 3113 Bellevue Ave, Cincinnati, OH 45219. E-mail: grohkn@ucmail.uc.edu
Journal of Voice, Vol xx, No xx, pp. xxx–xxx
0892-1997

https://doi.org/10.1016/j.jvoice.2023.07.006

protect voice-over actors while maintaining their viability as artists in the video game industry.

Behavioral voice intervention techniques are evidence-based standards for the prevention and rehabilitation of dysphonia. Intervention approaches can be dichotomized between direct and indirect techniques: indirect techniques most commonly refer to basic vocal hygiene education and didactic knowledge, while direct techniques are those that modify specific aspects of voice production to promote improved vocal efficiency.[11] Indirect techniques have been shown to provide some benefit to those at risk for voice disorders, including radio professionals, actors, singers, and teachers.[12–15] However, indirect techniques have largely been shown to be less effective than direct techniques, particularly for individuals with voice disorders.[11,16–20] Currently, there are no professional standard requisites for direct occupational voice training for voice-over actors. Though direct vocal training programs to minimize phonotrauma and dysphonia exist,[21–23] these "open throat" techniques were not specifically developed to preserve realism and authenticity, and may not be consistent with video game industry standards. As such, the development and validation of effective direct vocal techniques that are occupation-specific are needed for voice-over actors who perform in violent video games.

Vocal Combat Technique (VCT) training was developed by authors D.S. and K.R. to provide both indirect and direct behavioral techniques to voice actors. This training takes a three-pronged approach to include: i) vocal hygiene education, ii) vocal resonance and breath support techniques for healthy phonation, and iii) specialized voice training for yells, screams, and commands. Our previous pilot study demonstrated that self-reported symptoms (eg, Voice Handicap Index-10 [VHI-10]) and acoustic outcomes in five voice-over actors who completed violent video game voice-over recordings improved after they completed VCT training.[24] Although this study was promising, the small sample size of this study precluded the ability to perform inferential statistical analyses, and furthermore, the design did not include a control group for comparison.

Therefore, the purpose of the present study was to build upon our previous pilot investigation to understand whether VCT assists in mitigating vocal symptoms for voice-over actors while maintaining realism. To do so, we designed a prospective, double-blinded, randomized, controlled, experimental design, which directly compared the efficacy and realism of VCT to training with indirect voice techniques alone. We hypothesized that the experimental group that received VCT would report lower amounts of vocal handicap and a higher likelihood to return to performance while preserving vocal realism compared to the control group.

METHODS

This study was performed using a parallel-group randomized design. As such, we followed the CONSORT guidelines set forth to improve reporting and transparency of parallel-group randomized trials, including the 25-item checklist to ensure reproducibility.[25]

Participants

The Institutional Review Board approval was obtained through the University of Cincinnati for all recruitment and study procedures. Participants were recruited with the help of ACTRA, from the Toronto, Canada-based professional voice-over acting community, in response to an emailed survey. Criteria for enrollment included English-proficient voice-over actors, aged 18–65 years, who had participated in at least one AAA-level video game, indicating experience working on a high-caliber video game.[26] Participants were excluded if they: i) had any prior experience in similar vocal combat training, ii) self-reported history of vocal injury, and/or iii) were currently actively engaged in voice-over work one week prior to and/or throughout the duration of the study period. This third exclusion criterion was to ensure that there were no lingering effects of previous voice-over work, and that participants were vocally rested.

A total of 40 individuals responded to the inclusion/exclusion statements and were screened for eligibility; all met the inclusion criteria. We then used a random number generator to randomly select 24 eligible adults to enroll in the study. Only 24 were selected due to the proposed sample size estimation that resulted in 12 participants per group, which ensured statistical power. Selected participants then completed informed consent to formally enroll in the experimental portion of the study. Next, consented participants were randomized into either the VCT experimental group (n = 12) or the active treatment (vocal hygiene) control group (n = 12) in a 1:1 equal allocation. Participants were blinded to which exposure they received. No participants were lost to follow-up, and all participants' data were included in the statistical analyses. Please see Figure 1 for a flow diagram of enrollment, allocation, follow-up, and analysis.

Experimental exposure and control

Participants randomized to the experimental group attended a standard 3-hour VCT workshop led by author D.S. VCT is a combination of indirect and direct techniques, covering three target areas: 1) vocal hygiene: education regarding vocal conditioning to maximize stamina and expedite recovery after vocal exertion and strategies to reduce laryngeal irritation; 2) vocal foundations: direct physiologic and motor learning instruction to target respiratory-phonatory support, modification of phonatory onsets, and forward focus resonance techniques, to optimize laryngeal function; and 3) vocal effects: in-depth instruction on the production of aggressive vocalizations (eg, shouting, screaming, coughing, and growling) that utilizes supraglottic phonation techniques and reduces phonotraumatic behaviors. VCT training was provided online in a group Zoom session so that each participant received the

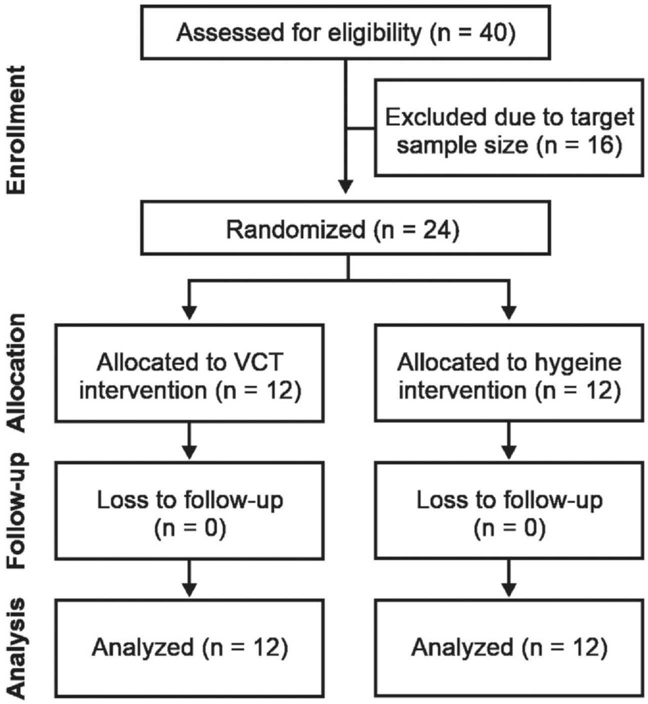

FIGURE 1. Flow diagram of enrollment, allocation, follow-up, and analysis. VCT, vocal combat technique.

same training and experience. The training was interactive, where participants practiced techniques and received feedback to shape their vocalizations into the desired targets. Participants were encouraged to use the techniques in and out of the session; however, no further training was provided beyond this single training session.

Participants randomized to the active treatment control group received only indirect techniques, consisting of a 45-minute educational vocal hygiene session. Information was presented through an online group Zoom session. This session included the same vocal hygiene information provided to the VCT group. Specific information included: i) education on behavioral modifications for reducing laryngeal irritants (reflux and smoking), ii) healthy habits to incorporate into their lifestyle (hydration and rest), and iii) warm-up and cool-down techniques for before and after the performance (eg, humming, straw phonation, and resonant voice). Once again, participants were encouraged to use these techniques outside of the session and in their daily lives and/or performances, but no further training or encouragement was provided beyond the single training session.

Study procedure

Each participant completed a demographics survey, including information on experience, professional voice use, and years of vocal training. They also completed the VHI-10[27] and a modified version of the Evaluation of the Ability to Sing Easily (m-EASE) Questionnaire. The original EASE is a 22-item survey rated on a Likert scale from 1 to 4. The rating scale assesses the degree of the problem with ratings of 1: "not at all," 2: "mildly," 3: "moderately," and 4: "extremely." We used a modified survey derived from the original EASE that was developed to assess singing voice.[28] For example, the prompt *"I am having difficulty singing softly"* was changed to *"I am having difficulty speaking softly."*

Two to 3 days after their VCT training or control exposure, each participant then underwent a 1-hour long simulated video game recording session (SVRS). The timeframe of at least 48 hours following VCT training was chosen because of the well-known industry standard and the expectation that actors should be recovered within 48 hours (about 2 days) and be ready to work again. The SVRS was completed in a real voice-over recording studio that was set up exactly how a real video game recording session would occur. The participants stood for all target utterances. Although rest breaks were allowed, no participant chose to take one.

During each SVRS, an industry professional with over 25 years of voice-over directing experience led the session. The director was blind to the participant's group assignment, and thus, all participants received the same instructions. No additional instruction, feedback, or encouragement was provided to use VCT or hygiene techniques during the session. Participants were asked to perform vocalizations commonly required for video games. These included, but were not limited to, grunts, growls, yelling commands, and screaming. All participants completed all targets, which included a total of 284 utterances. The director was instructed to subjectively rate the realism and acceptability of the targets based on video game industry standards. Performances were rated as "satisfactory" or "unsatisfactory" for each of the 284 utterances.

Immediately following the SVRS, the Vocal Tract Discomfort Scale (VTDS) was completed.[29,30] The VTDS is a Likert-based assessment that asks about the frequency (never-to-always) and the severity (none-to-extreme) of vocal tract discomfort. Ratings are made across eight different sensations: burning, tight, dry, aching, tickling, sore, irritable, and lump in the throat. Higher ratings indicate a greater amount of discomfort.

Then, 48 hours after the SVRS, participants were sent and remotely completed a postperformance VHI-10 and m-EASE assessment. Additionally, participants completed a three-question survey on a Likert scale. Question 1 was, *"How much difficulty did you have returning to your daily vocal demands?"* with possible responses: None, A mild amount, A moderate amount, and An extreme amount. Question 2 was, *"How likely are you to book another recording sessions today?"* with possible responses: Very likely, Somewhat likely, and Not likely. Question 3 was, *"How confident are you to safely record another voice-over session?"* with possible responses: Very confident, Confident, Somewhat confident, and Not confident.

Outcome measures

The VHI-10 was used as the primary outcome measure for this randomized controlled trial. All other questionnaires and ratings were considered secondary outcomes. For both the VHI-10 and the m-EASE, the difference between pre-performance/postperformance was calculated for each person and then compared across the groups. The measures of the VTDS, three-question survey, and voice director ratings of realism were only measured at one timepoint, so were directly compared between groups.

Statistical analysis

Our previous pilot investigation of pre/post within-participant comparison of VCT intervention showed a large effect size improvement (Cohen's $d = 1.95$) for the VHI-10.[24] Using a change in VHI-10 score as the primary outcome measure, we were able to calculate the appropriate sample size for the present study that included the standard assumptions of $\alpha = 0.05$ and an expected large effect size ($d = 1.95$) difference between the control and experimental groups. A sample of 24 participants, with 12 participants in each group, allowed for a power of 0.95, mitigating the likelihood of both type I and type II errors in our primary analysis.

Shapiro-Wilk tests were completed to assess the parametric distribution of the continuous variables, finding that all continuous variables were not normally distributed (all $P < 0.05$). Thus, nonparametric Wilcoxon rank sum tests were employed to determine differences between groups for these measures. Fisher's exact tests were used to evaluate categorical variables (eg, participant demographics). *Post hoc* analysis was completed on the VHI-10 scores to identify outliers and investigate their differences across all assessments. A cutoff criterion for the interquartile range was used to identify outliers using the equation: VHI-10 > 1.5 × interquartile range + Q3 VHI-10,[31] resulting in an outlier criterion of change in VHI-10 > 7.4. Significance was set to $P < 0.05$ for all analyses. Statistical analysis and figures were created with "rStudio," Foundation for Statistical Computing, Vienna, Austria (URL: https://www.R-project.org/; packages "tidyverse" and "stringer").

RESULTS

Participant characteristics

All 24 study participants completed all tasks within the study. Participants included 14 females and 10 males, aged 25–64 years (mean = 41.6 years and standard deviation [SD] = 10.8 years). Baseline demographics, experience, training, and reported voice use did not differ between the experimental and the active treatment control group ($P > 0.05$; see Table 1). Reported VHI-10 and m-EASE scores before the SVRS were also similar between groups at baseline ($P > 0.05$).

Primary and secondary outcomes

There was a larger increase in VHI-10 noted in the active treatment control group compared to the VCT group with a median difference increase of 4 (95% confidence interval [CI] 1,16; $P = 0.004$). There was also a larger increase in m-EASE noted in the active treatment control group compared to the VCT group with a median difference increase of 9 (95% CI 1.5, 20; $P = 0.049$).

TABLE 1.
Baseline Demographics, Experience, Voice Use, and Vocal Training Between the VCT and the Active Treatment Control Groups

	VCT group	Control group	*P* value
Age	39 yrs (11)	43 yrs (11)	0.15
Gender			0.68
Male	6 (50%)	4 (33%)	
Female	6 (50%)	8 (67%)	
Race*			0.35
White	9 (75%)	6 (50%)	
Black	2 (17%)	1 (8%)	
Asian	1 (8%)	3 (25%)	
Mixed	0 (0%)	2 (17%)	
Acting experience*	19 yrs (10)	21 yrs (11)	0.55
Weekly voice use*	7 h (5, 13)	10 h (3, 13)	0.28
Voice training	3 yrs (2, 11)	2 yrs (1, 8)	0.27
Baseline VHI-10	4 (0, 9)	0.5 (0, 3)	0.70
Baseline m-EASE	30 (26, 35)	26 (25, 31)	0.24

Abbreviations: VCT, vocal combat technique; VHI-10, voice handicap index-10; m-EASE, modified evaluation of ability to sing easily; hrs, hours; yrs, years.
Reported as mean (Standard Deviation), median (Q1, Q3), and n (%) for parametric, nonparametric, and categorical variables.
* Self-reported.

VTDS was only collected at one timepoint (post-SVRS), and the differences were assessed between groups. VTDS scores for both frequency and severity showed no statistical difference between the two groups (Table 2, Figure 2). Realism also did not differ significantly between groups. Please see Figure 3 for boxplots of realism ratings for different voice-over tasks.

In the postperformance survey, participants in the VCT group more frequently stated that they were "confident" or "very confident" to safely perform another video game voice-over session compared to control participants ($n = 11$ versus $n = 4$, $P = 0.01$). The experimental group more frequently stated that they were more "likely" or "very likely"

TABLE 2.
Changes in VHI-10, m-EASE, and the Values for VTDS. Positive Values Indicate More Voice Symptoms

	VCT median (Q1, Q3)	Control median (Q1, Q3)	Median difference (95% CI)
Change in VHI-10	−2 (−3,0)	1 (0,14)	4 (1, 16)*
Change in m-EASE	6 (1,11)	16 (3, 24)	9 (1.5, 20)†
VTDS Severity	15 (6,19)	9 (6,17)	−1 (−10, 7)
VTDS Frequency	11 (5,18)	7 (4, 10)	−2 (−11, 4)

Median difference, VCT versus control, derived from the coefficient estimate from the wilcoxon rank sum test.
VHI-10, voice handicap index-10; m-EASE, modified evaluation of ability to sing easily; VTDS, vocal tract discomfort scale; CI, confidence interval; hrs, hours; yrs, years.
* $P < 0.05$.
† $P < 0.01$.

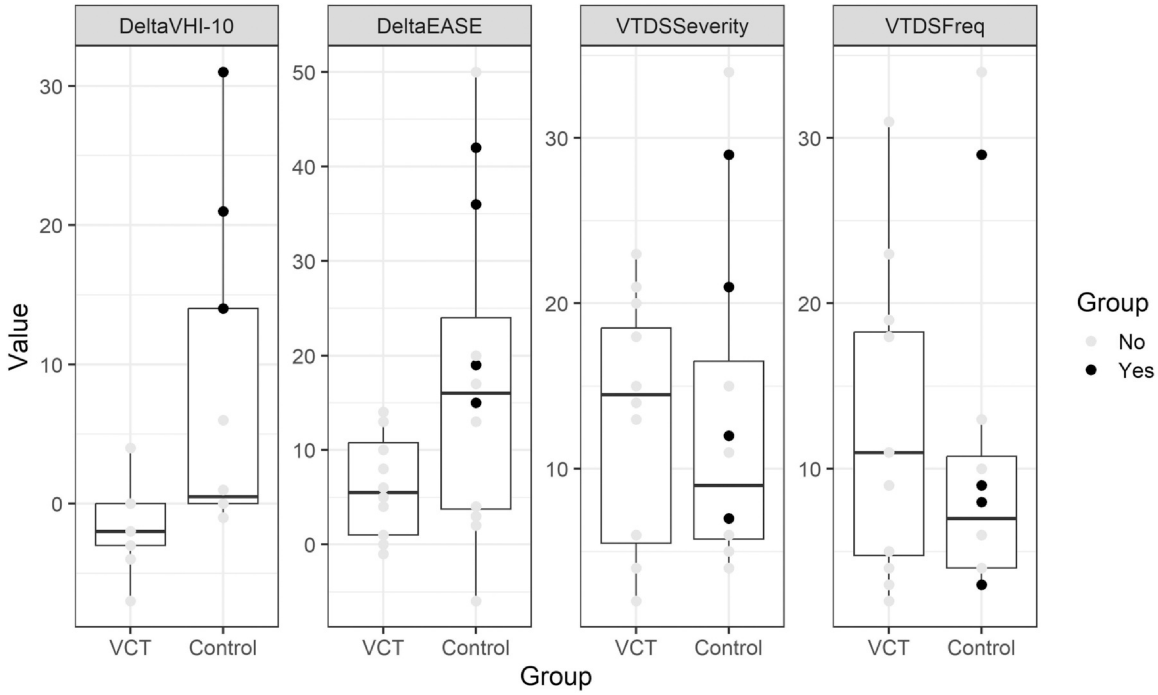

FIGURE 2. The scores of the patient reported outcomes measures from the postperformance survey. Outliers, as defined in the post hoc analysis, are shown by dot transparency as noted in the legend by outlier group (yes, outlier; no, not an outlier). VCT, vocal combat technique; DeltaVHI-10, change in Voice Handicap Index-10, positive scores indicate worsening vocal symptoms; DeltaEASE, change in the modified version of the Evaluation of Ability to Sing Easily, positive scores indicate more difficulty with vocal performance; VTDS Severity, Vocal Tract Discomfort Scale—severity domain; VTDSFreq, Vocal Tract Discomfort Scale—frequency domain.

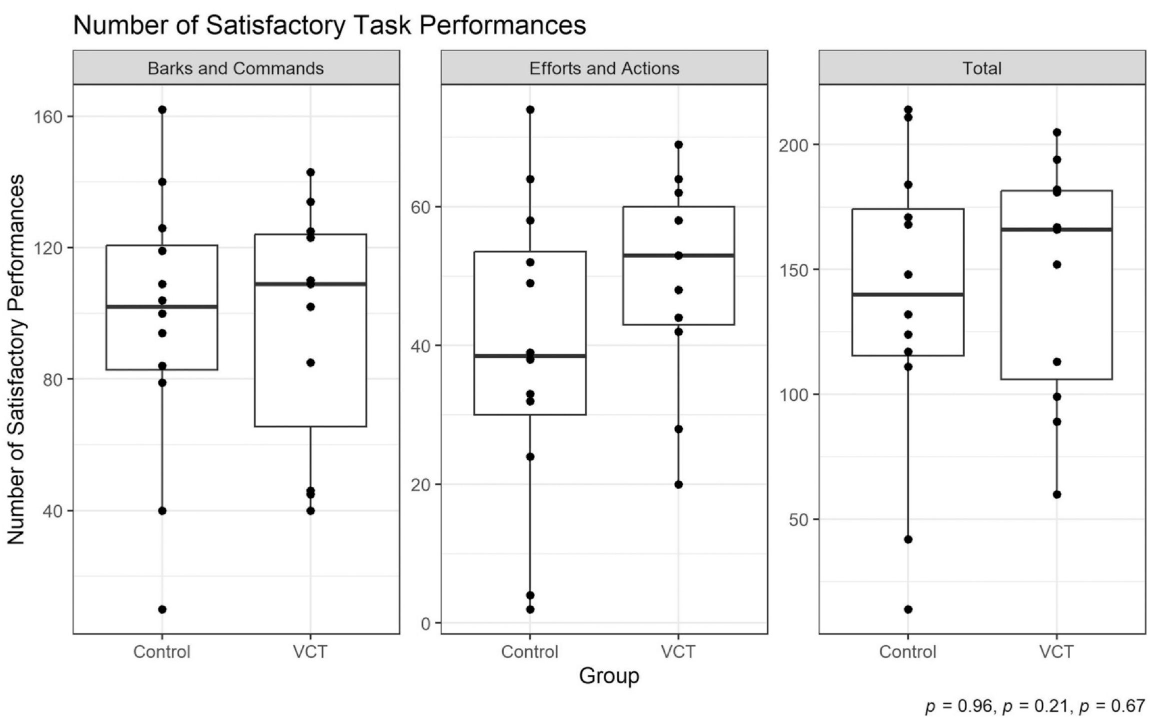

FIGURE 3. The number of satisfactory performances as rated by an experienced video game director. Performance was divided into a series of barks and commands, and efforts and actions. Total is the sum of the two subcategories. Significance values (*P* values) are for comparisons between the vocal combat technique (VCT) and the active treatment control groups for each series, respectively.

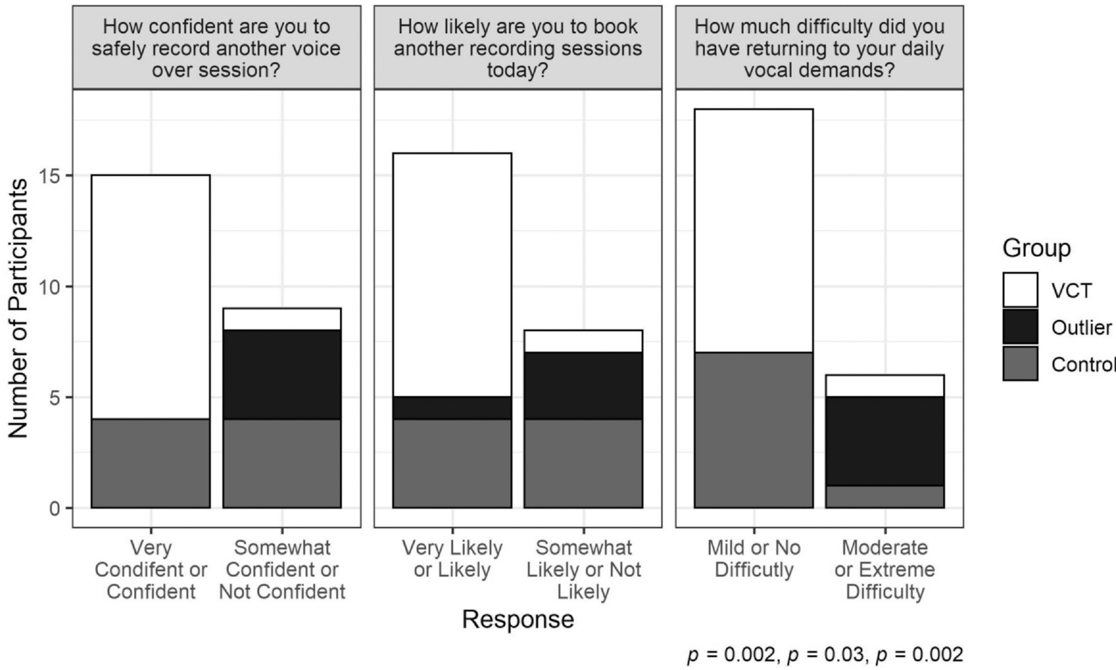

FIGURE 4. Responses to the three postperformance survey questions including post hoc analysis that identified outlying participants. Shading according to exposure as noted in the legend. Significance values (P values) are for comparisons between the outliers and all other responses from the vocal combat technique (VCT) and the active treatment control groups.

to register for another recording session today ($n = 11$ versus $n = 5$, $P = 0.02$) compared to the active treatment control group. There was no difference in reported ease to return to normal voice use ($n = 11$ versus $n = 7$, $P = 0.145$).

Outlier analysis

The post hoc analysis identified a group of four individuals in the active treatment control group, who were found to be outliers by their change in VHI-10 (defined as a change in VHI-10 > 7.4 [1.75 + 3.75 × 1.5]. These outliers had a mean (± SD) increase in VHI-10 score of 20 (± 8) and a mean increase in m-EASE of 28 (± 13). This outlier group was also significantly less likely to sign up for another voice-over session, easily return to normal voice use, and to have the confidence to safely perform voice-over sessions (Figure 4). The outlier group had no significant differences in age, gender, race, voice-over experience, duration of vocal training, amount of routine voice use, baseline VHI-10, or baseline m-EASE compared to both the remainder of the active treatment control group and VCT group.

DISCUSSION

VCT is a promising training that combines indirect and direct behavioral techniques for vocal actors. This randomized controlled trial was the first to evaluate whether VCT training would mitigate negative effects for voice actors performing intensive video game recordings compared to an active treatment control group of voice actors who received vocal hygiene training alone. Our hypotheses were supported when voice actors in the VCT group showed significantly mitigated vocal handicap and dysphonia symptoms without difference to the overall realism of their vocal performance compared to the control group.

This study provides evidence for VCT as a technique to maintain healthy vocal function for vocal performers. The VCT group showed significantly smaller amounts of change in pre-/post-SVRS scores for both the VHI-10 and m-EASE compared to the active treatment control group. That is, the control group increased a median difference of four and nine over the VCT group for both questionnaires, respectively. For our primary outcome measure of the VHI-10, the median difference of four between the changes within groups corresponds to a clinically meaningful difference in the degree of handicap and dysphonia symptoms, which has previously been defined as a range of 4–6 points for those with clinical voice disorders.[32,33] However, this difference did not meet clinical significance when examining change within each group. The VCT group showed a median improvement (decrease) of two points on the VHI-10, whereas the active treatment control group showed an increase in score by 1 point (indicating a slightly higher degree of handicap).

Our pilot study of five voice-over actors revealed results comparable to the current study, where the average decrease across sessions with and without VCT training on the VHI-10 was also four points and the decrease across sessions for m-EASE was eight points.[24] Interestingly, this comparison demonstrates that receiving vocal hygiene education resulted in the same impairment outcomes as when no intervention was provided in the pilot study. This may have occurred for several reasons. For one, the potential impact of vocal hygiene education is relative to the degree of vocal hygiene impairment that an individual is experiencing at baseline.[15] As such, voice actors with healthier habits may have less to gain from this type of intervention. Additionally, it should be noted that, even for those who may find vocal hygiene adjustments advantageous based on their current lifestyle and habits, changes in behaviors can come in small increments that may take time to reach significance,[34] and therefore, resulting physical correlates are unlikely to be perceived immediately. It is then possible that the delivery of this information two days prior to the SVRS did not allow adequate time for implementation and, therefore, impact. However, prior research has also shown that vocal hygiene education is not effective at mitigating phonotrauma when vocally demanding tasks are necessary (eg with stage performance).[13] This suggests that this type of intervention alone is insufficient for protecting against intense and aggressive laryngeal demands and behaviors. Taken together, this highlights the benefits of a direct training approach like VCT for this population as it did result in a statistically significant reduction in perceived vocal symptoms and handicap across both the VHI-10 and m-EASE.

Consistent with these results, those with VCT training were significantly more likely to report a higher likelihood to return to work and demonstrate a high level of confidence in safely recording subsequent voice-over sessions. This is important because it directly addresses an issue identified by more than a quarter of the surveyed ACTRA members who reported that they considered turning down a session for fear of the impact that it would have on their voice or work that it would cause them to lose.[10] It is also of note that the industry expects voice-over artists to be fully recovered by 48 hours after their last recording session. However, as the ACTRA survey revealed, approximately half of voice-over artists do not feel fully recovered for two or more days following these sessions. With industry demand for performers increasing,[4] utilization of VCT could mean increasingly productive, profitable, and lengthy careers for voice-over artists attempting to make a living in this field.

Importantly, realism was maintained for the VCT group compared to the active treatment control group. Both groups showed performance accuracy ~50%–60%. These accuracies were similar to our previous findings in which five voice actors showed an average performance accuracy of 47% prior to receiving VCT training and showed an increase to 73% after receiving VCT training.[24] In the present study, realism for the VCT group could have been maintained due to the specialized training aspects of the program. Many of the established techniques for the healthy production of aggressive sounds were designed for stage performance and focused on an open vocal tract with only the minimum amount of turbulence or distortion deemed necessary for the audience to understand the emotion/intent behind the vocalization.[21–23] These techniques do not always translate well into the video game recording booth, as is evidenced by our study: our participants had an average of 20 years of acting experience and two to three years of vocal training, yet the active treatment control group admitted to having little confidence in their ability to safely perform the tasks often required of them. VCT is unique because it was designed with the video game industry standards in mind and provided training in alternative methods of creating the target sounds without compromising their quality or believability/realism. The success of this method in achieving this is evidenced by our results; VCT was not only able to mitigate perceived vocal handicap, but it did so without compromising the quality or accuracy of the sounds, which is paramount for the viability of such a technique. In real-world applications, performers are only required to do as many repetitions, as is necessary to reach a realistic target. Thus, increasing the realism of performances acts to reduce vocal load during a session and may also prove to be beneficial for actors to maintain vocal health.

However, there were no mitigating immediate reports of vocal tract discomfort and laryngeal symptoms via the VTDS. As VCT utilizes turbulence and constriction of several supraglottic structures to add distortion to the sound production (pharyngeal walls, superior arytenoid cartilages, aryepiglottic folds, epiglottis, etc), it is possible that the aperiodic/chaotic vibration of these structures may also cause some degree of discomfort. We can endeavor to understand this further with research into the specific physiologic maneuvers of VCT (eg via visualization on flexible laryngoscopy) and by acquiring further objective data about the types of sensations experienced and, in particular, their durations. Higher frequencies of vocal discomfort are found in those with self-reported voice problems and dysphonia compared to those without.[35,36] The way in which the VTDS was utilized in our study only provided a snapshot of information about the frequency of the types of vocal tract discomfort that occurred during the recording session and immediately following. Application of this questionnaire a second or third time at follow-up intervals (eg mid-week and again at one week post-SVRS) may have provided more information about the potentially enduring nature of some of the experienced sensations. It would be valuable to quantify sensations and discomforts in the longer term for both groups to understand when each group was able to return to their baseline level of comfort and if the groups were, in fact, equal.

Through our post hoc analysis, we identified four participants from the active treatment control group who exhibited more severe vocal handicap impairments, as defined by a change > 7.4 on the VHI-10. When examined further, these participants also showed consistently high scores in the m-EASE and the VTDS severity assessment, and reported being more likely to have difficulty in returning to their baseline vocal quality, less confidence in their vocal ability, and lower likelihood to sign up for another voice-over session compared to all other participants in the study. Because these individuals represented one-third of the control sample, they were responsible for the large variance in the control data and contributed to the significant differences observed between the two groups. Despite these trends in the outlier data, there was no single demographic characteristic that sets the outliers apart from the remainder of the control group nor from the VCT group. Therefore, more work is needed to understand the best way to identify voice-over actors at risk for more adverse responses to performing aggressive vocal behaviors. Previous work has identified risk factors of laryngopharyngeal reflux,[37–39] environmental allergies,[40] smoking history,[41] and systemic/surface hydration,[42] which were not assessed within the present study. The next steps in this line of work should consider that these are possible risk factors associated with vocal fatigue and the manifestation of dysphonia. If actors with these increased risk factors could be identified through simple surveys of characteristics and exposures, they could then be targeted for preventative techniques to ameliorate their vocal symptoms and decrease their experiences of postperformance dysphonia.

Limitations and future directions

The current study primarily used self-reported outcome measures to understand the impact that VCT could have on vocal handicap and vocal symptoms. However, no objective measures or formalized instrumental assessments were completed, limiting our interpretation of the results. Further work could include high-speed laryngoscopy and/or respiratory plethysmography to better capture physiological changes associated with VCT training. High-speed laryngoscopy has an advantage over standard stroboscopy when it comes to capturing aperiodic vocalizations (eg, grunts and growls) because of the inability of stroboscopy to effectively synchronize images with short (< 3 seconds) and/or inconsistent acoustic waveforms. The next steps could also include acoustic parameters as possible markers of dysphonia, such as cepstral peak prominence, which has been shown to be related to overall vocal quality and auditory-perceptual ratings of dysphonia.[43] Furthermore, acoustic markers may be collected over time and provide temporal information about longer recovery times. One reason that we did not pursue acoustical analysis in this study was because of the challenges related to completing this study during the COVID-19 pandemic. Future studies should include a controlled environment to collect high-quality acoustic recordings for acoustic analysis.

Of the questionnaires used in this study, the VHI-10 was the most appropriate for our primary outcome measure, as it has been used as a vocal outcome across several other voice studies[44–47] and has undergone rigid psychometric testing.[27] The original EASE, although validated in singers,[28] was adapted for the purposes of the present study so that it would be relevant to our study population. Our modifications from the word “sing” to “speak” limited our ability to compare our findings of the m-EASE to other studies and across other populations. Since the completion of our study, the Evaluation of Ability to Voice Easily (EAVE)[48] was developed as a 14-question survey to assess physical symptoms during speaking. This fully validated questionnaire may be appropriate for further investigations into vocal functions in voice-over actors.

Our study paradigm employed a single session of VCT prior to SVRS. It must also be taken into consideration that one training session may not be adequate to fully reveal the potential benefits of the technique. It is possible that some performers may have more difficulty mastering the technique immediately after introduction to it and less than 48 hours to practice before attempting real-world application. With increased practice and instruction in the methods, actors’ accuracy with producing these sounds may increase and, as such, so might their comfort. At present, formalized VCT workshops provide instruction for 6 hours over the course of 1-week (for industry professionals) or 30 hours over a 10-week program (for students just beginning their training). Current training programs are longer than the single session 3-hour training provided here. Determining the adequate dosing and duration of VCT training, as well as whether additional 1:1 training and/or refresher training are needed, is important to provide the right foundational support for voice actors and should be investigated further.

Finally, we could not identify any specific traits that were different within our outlier group compared to other members of either group. Perhaps more objective measures, patient-reported ratings of vocal effort, or more specific demographic information (such as how long it has been since their last AAA-level video game voice-over) could have helped us to further distinguish these individuals from the group. Another possibility is that self-reports of vocal load may be inaccurate due to known difficulties in individuals quantifying and reporting vocal demands and speaking durations.[49] Recent advancements in ambulatory monitoring of vocal dose[50,51] could be advantageous to this population. Then, we could work to identify people at risk for more pronounced vocal symptoms and direct them toward training resources (such as VCT) to help them mitigate the effects of the demanding vocal recording sessions that the video game industry requires.

CONCLUSIONS

This randomized controlled study provides strong evidence that VCT, a combination of direct and indirect training techniques, can mitigate postperformance vocal symptoms and preserve the realism required for performers in the video game industry. Furthermore, VCT improved participants’ perceived ability to return to work and complete subsequent voice-over performances, providing a way for them to continue to engage in the profession. Future study is needed to identify risk factors for postperformance dysphonia symptoms to best target VCT training to individuals who will benefit from them.

DECLARATION OF COMPETING INTEREST

D′Arcy Smith is the developer of Vocal Combat Technique. Katelyn Reid has been paid to provide lectures on the vocal mechanism and vocal health at Vocal Combat Technique

workshops. All other authors have no declarations of interest for this work.

ACKNOWLEDGMENTS

The authors thank Alliance of Canadian Cinema, Television and Radio Artists (ACTRA) for their assistance in recruitment of experienced video game voice-over actors and Kim Hurdon from Kim Hurdon Casting for her work as voice director during the sessions and evaluator of quality in performance.

REFERENCES

1. M. Luenendonk. The Gaming Industry—An Introduction. Cleverism 2019. Available at: https://www.cleverism.com/gaming-industry-introduction/. Accessed February 25, 2023.
2. Saltzman. Video games are bigger business than ever, topping movies and music combined. Published June 10, 2021. Available at: https://www.usatoday.com/videos/tech/2021/06/10/e-3–2021-video-games-big-business-topping-film-and-music-combined/7637695002/. Accessed March 26, 2023.
3. Clement J. Most popular video game genres worldwide 2022, by age group. Published January 31, 2023. Available at: https://www.statista.com/statistics/1263585/top-video-game-genres-worldwide-by-age/. Accessed March 26, 2023.
4. Hern, A.Game (voice)over: actors turn to video game work during pandemic. The Guardian. Published 2021. Available at: https://www.theguardian.com/games/2021/aug/15/actors-turn-to-video-game-work-during-pandemic. Accessed March 26, 2023.
5. Hunter EJ, Cantor-Cutiva LC, van Leer E, et al. Toward a consensus description of vocal effort, vocal load, vocal loading, and vocal fatigue. *J Speech Lang Hear Res.* 2020;63:509–532.
6. Pestana PM, Vaz-Freitas S, Manso MC. Prevalence of voice disorders in singers: systematic review and meta-analysis. *J Voice.* 2017;31:722–727.
7. Roy N, Merrill RM, Gray SD, et al. Voice disorders in the general population: prevalence, risk factors, and occupational impact. *The Laryngoscope.* 2005;115:1988–1995.
8. Guss J, Sadoughi B, Benson B, et al. Dysphonia in performers: toward a clinical definition of laryngology of the performing voice. *J Voice.* 2014;28:349–355.
9. Cazden J. Screaming for attention: the vocal demands of actors in violent interactive games. *J Voice.* 2017;31:1–2.
10. ACTRA. Vocal Stress in Video Game Voiceover Survey. Published August 4, 2022. Available at: https://www.actratoronto.com/vocal-stress-in-videogame-voiceover-survey/. Accessed March 26, 2023.
11. Carding PN, Horsley IA, Docherty GJ. A study of the effectiveness of voice therapy in the treatment of 45 patients with nonorganic dysphonia. *J Voice.* 1999;13:72–104.
12. Chan RWK. Does the voice improve with vocal hygiene education? A study of some instrumental voice measures in a group of kindergarten teachers. *J Voice.* 1994;8:279–291.
13. Rangarathnam B, Paramby T, McCullough GH. "Prologues to a bad voice": effect of vocal hygiene knowledge and training on voice quality following stage performance. *J Voice.* 2018;32:300–306.
14. Roy N, Weinrich B, Gray SD, et al. Voice amplification versus vocal hygiene instruction for teachers with voice disorders: a treatment outcomes study. *J Speech Lang Hear Res.* 2002;45:625–638.
15. Timmermans B, De Bodt M, Wuyts F, et al. Vocal hygiene in radio students and in radio professionals. *Logoped Phoniatr Vocol.* 2003;28:127–132.
16. Bassiouny S. Efficacy of the accent method of voice therapy. *Folia Phoniatr Logop.* 1998;50:146–164.
17. Behrman A, Rutledge J, Hembree A, et al. Vocal hygiene education, voice production therapy, and the role of patient adherence: a treatment effectiveness study in women with phonotrauma. *J Speech Lang Hear Res.* 2008;51:350–366.
18. Gartner-Schmidt JL, Roth DF, Zullo TG, Rosen CA. Quantifying component parts of indirect and direct voice therapy related to different voice disorders. *J Voice.* 2013;27:210–216.
19. Holmberg EB, Hillman RE, Hammarberg B, et al. Efficacy of a behaviorally based voice therapy protocol for vocal nodules. *J Voice.* 2001;15:395–412.
20. Speyer R. Effects of voice therapy: a systematic review. *J Voice.* 2008;22:565–580.
21. Roy N, Ryker KS, Bless DM. Vocal violence in actors: an investigation into its acoustic consequences and the effects of hygienic laryngeal release training. *J Voice.* 2000;14:215–230.
22. Snyder T. Retract, then attack. *Fight Master.* 2012;34:18–19.
23. Raphael BN. The sounds of violence: vocal training in stage combat. *Theatre Top.* 1991;1:73–86.
24. Reid K, McKenna VS, Smith D. Mitigating dysphonia, pain, and vocal handicap after violent video game voice overs: a pilot investigation into vocal combat technique training. *J Voice.* 2022. https://doi.org/10.1016/j.jvoice.2022.06.015.
25. Schulz KF, Altman DG, Moher D. for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340:c332. https://doi.org/10.1136/bmj.c332.
26. Steinberg S. *Videogame Marketing and PR. 1: Playing to Win: [Top the Charts; Make Headlines; Earn Better Reviews].* iUniverse; 2007.
27. Rosen CA, Lee AS, Osborne J, et al. Development and validation of the Voice Handicap Index-10. *Laryngoscope.* 2004;114:1549–1556.
28. Phyland DJ, Pallant JF, Thibeault SL. Measuring vocal function in professional music theater singers: construct validation of the Evaluation of the Ability to Sing Easily (EASE). *Folia Phoniatr Logop.* 2014;66:100–108.
29. Mathieson L. Vocal tract discomfort for hyperfunctional dysphonia. *Voice.* 1993;2:40–48.
30. Mathieson L, Hirani SP, Epstein R, Baken RJ, Wood G, Rubin JS. Laryngeal manual therapy: a preliminary study to examine its treatment effects in the management of muscle tension dysphonia. *J Voice.* 2009;23:353–366.
31. Tukey JW. *Exploratory Data Analysis: "Box and Whisker Plots".* Reading, MA: Addison-Wesley; 1977.
32. Misono S, Yueh B, Stockness AN, et al. Minimal important difference in Voice Handicap Index–10. *JAMA Otolaryngol Neck Surg.* 2017;143:1098.
33. Young VN, Jeong K, Rothenberger SD, et al. Minimal clinically important difference of Voice Handicap Index-10 in vocal fold paralysis. *Laryngoscope.* 2018;128:1419–1424.
34. Broaddus-Lawrence PL, Treole K, McCabe RB. The effects of preventive vocal hygiene education on the vocal hygiene habits and perceptual vocal characteristics of training singers. *J Voice.* 2000;14:58–71.
35. Lopes LW, de Oliveira Florencio V, Silva POC. Vocal Tract Discomfort Scale (VTDS) and Voice Symptom Scale (VoiSS) in the evaluation of patients with voice disorders. *J Voice.* 2019;33:381.e23–381.e32.
36. Rodrigues G, Zambon F, Mathieson L. Vocal tract discomfort in teachers: its relationship to self-reported voice disorders. *J Voice.* 2013;27:473–480.
37. Ayazi S, Pearson J, Hashemi M. Gastroesophageal reflux and voice changes: objective assessment of voice quality and impact of antireflux therapy. *J Clin Gastroenterol.* 2012;46:119–123.
38. Lechien JR, Saussez S, Harmegnies B. Laryngopharyngeal reflux and voice disorders: a multifactorial model of etiology and pathophysiology. *J Voice.* 2017;31:733–752.
39. Spantideas N, Drosou E, Karatsis A. Voice disorders in the general greek population and in patients with laryngopharyngeal reflux. prevalence and risk factors. *J Voice.* 2015;29:389.e27–389.e32.
40. Hamdan AL, Abi Zeid Daou C, Karam M. Prevalence of allergy in patients with primary dysphonia. *Indian J Otolaryngol Head Neck*

Surg. 2022;75(2):692–696. https://doi.org/10.1007/s12070-022-03403-8.Published online December 30.

41. Byeon H. The association between lifetime cigarette smoking and dysphonia in the Korean general population: findings from a national survey. *PeerJ.* 2015;3:e912. https://doi.org/10.7717/peerj.912.
42. Alves M, Krüger E, Pillay B. The effect of hydration on voice quality in adults: a systematic review. *J Voice.* 2019;33:125.e13–125.e28.
43. Awan SN, Roy N, Dromey C. Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model. *Clin Linguist Phon.* 2009;23:825–841.
44. Al Afif A, Rigby MH, MacKay C, et al. Injection laryngoplasty during transoral laser microsurgery for early glottic cancer: a randomized controlled trial. *J Otolaryngol - Head Neck Surg.* 2022;51:12.
45. Dhaliwal SS, Doyle PC, Failla S. Role of voice rest following laser resection of vocal fold lesions: a randomized controlled trial. *The Laryngoscope.* 2020;130:1750–1755.
46. Lifante JC, McGill J, Murry T. A prospective, randomized trial of nerve monitoring of the external branch of the superior laryngeal nerve during thyroidectomy under local/regional anesthesia and IV sedation. *Surgery.* 2009;146:1167–1173.
47. Lin FC, Chien HY, Chen SH. Voice therapy for benign voice disorders in the elderly: a randomized controlled trial comparing telepractice and conventional face-to-face therapy. *J Speech Lang Hear Res.* 2020;63:2132–2140.
48. Phyland D, Pallant J, Free N. On the EAVE: a new scale for self-evaluation of the physical aspects of speaking voice function. *Presented at: IALP Voice Composium.* 2021.
49. Mehta DD, Cheyne HA, Wehner A. Accuracy of self-reported estimates of daily voice use in adults with normal and disordered voices. *Am J Speech Lang Pathol.* 2016;25:634–641.
50. Mehta DD, Zañartu M, Feng SW. Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *IEEE Trans Biomed Eng.* 2012;59:3090–3096.
51. Van Stan JH, Gustafsson J, Schalling E. Direct comparison of three commercially available devices for voice ambulatory monitoring and biofeedback. *Perspect Voice Voice Disord.* 2014;24:80–86.