Texas Christian University

Davies School of Communication Sciences & Disorders

The Strength of Association Between Self-Perceived Voice Handicap and Laryngeal Function Study Measurements

Master's Thesis

By

Emma Snodgrass

Fort Worth, Texas

2022

The Strength of Association Between Self-Perceived Voice Handicap and Laryngeal Function Study Measurements

A Thesis for the Degree

Master of Science

By

Emma Snodgrass

Thesis Approved By:

Christopher Watts
Dr. Christopher R. Watts, Ph.D., Thesis Advisor

Anthony DiLollo
Dr. Anthony DiLollo, Ph.D., CCC-SLP, Committee Member

Zoe Thijs
Dr. Zoë Thijs, Ph. D., Committee Member

Yan Zhang
Dr. Yan Zhang, Ph.D., L.Ac., Committee Member

Emily Lund
Dr. Emily Lund, Ph.D., Associate Dean of Research
Harris College of Nursing & Health Sciences

ACKNOWLEDGEMENTS

I would first like to express my deepest appreciation for Dr. Christopher Watts, my thesis advisor. Dr. Watts took me on as a “last minute” thesis student and supported me throughout the entire process. Having little experience with statistics and research, I am incredibly grateful for Dr. Watts and his patience as he guided me, encouraged me, and advised me in this process. I also would like to extend my sincere thanks to Dr. Zoë Thijs for agreeing to serve on my committee, her assistance throughout the process, and her willingness to meet with me to discuss various aspects of the study. Dr. Thijs took time out of her days to help me particularly with SPSS – from formatting data to analysis, I would not have been able to complete the methodology and analysis of this study without her guidance and patience.

I am also grateful for Dr. Yan Zhang and Dr. Anthony DiLollo for serving on my committee. Dr. Zhang provided valuable knowledge in the realm of statistics and was ever-patient in helping me wrap my head around methodology during my prospectus presentation. Dr. DiLollo was kind and helpful throughout the process, and his message of encouragement in the final week of writing my manuscript did not go unnoticed.

Thanks should also go to my friends and family for their support and encouragement. My dog, Autumn, was also my companion throughout the writing process.

I’d also like to acknowledge Kristie Knickerbocker of a Tempo Voice Center for providing the data utilized within this study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I: INTRODUCTION

1.1. Overview of the Study

Within the realm of voice disorders, historically research has found no significant correlation between clinical measures of voice and patient-perceived voice handicap (Wheeler, Collins, & Sapienza, 2006; Hsiung, Pai, & Wang, 2002). In recent years, however, there have been some studies to find significant, but weak correlations between the two (Karlsen, Sandvik, Heimdal, & Aarstad, 2020; Lopes, da Silva, Simoes, Evangelista, Silva, Almeida, & de Lima-Silva, 2017). Considering previous findings and limitations within those studies, this study aims to further expand the clinical understanding of the relationship between these measures. In particular, this study examines the correlation between acoustic (Acoustic Voice Quality Index-AVQI), aerodynamic (maximum phonation time – MPT), and auditory-perceptual (Consensus Auditory-Perceptual Evaluation of Voice – CAPE-V) measures of voice to patient-perceived voice handicap (Voice Handicap Index-30 and Voice Handicap Index-10 – VHI-30 and VHI-10). This study also examines the effect of disorder type on these measurements.

1. 2. Characteristics of Voice Disorders

Voice disorders result from deviations in vocal quality as compared to vocal norms defined by demographics such as age, sex, body type, and daily communicative or performance requirements. Typically, voice disorders are categorized as organic or functional. Organic voice disorders result from structural, systemic, or neurological changes to the mechanisms necessary for phonation. Functional voice disorders are those related to an imbalance or dysregulation of laryngeal and respiratory muscles, resulting in vocal behaviors that may or may not cause structural changes to the vocal folds. Whether the voice disorder is labeled functional or organic can substantially impact treatment options and recommendations provided to the patient. When considering the classification of voice disorders, it is important to remember that voice disorders are not limited to a single etiology, and that organic and functional voice disorders can co-occur (Watts & Awan, 2019). While the majority of voice disorders last less than four weeks in duration, the functional impact is significant. In a study performed by Roy et al. (2005), 7.2% of employed - reported missing at least one day of work in a year due to their voice disorders (Roy, Merrill, Gray, & Smith, 2005).

1.3. Prevalence and Incidence of Voice Disorders

The prevalence of dysphonia in the general population of adults is at or near 1% with the highest occurrence in adults 70 years of age or older; additionally, almost 30% of the adult population has experienced a voice disorder during their lifetime (Roy, Merrill, Gray, & Smith, 2005). In regard to what is known about the impact of gender and age on the prevalence and incidence of voice disorders, there appears to be a significant amount of intersectionality. For example, Martins et al. (2016) found a predominance of voice disorders in adults ages 20 to 60 years, with the majority being female (61%). However, males dominated the pediatric population in children one to 12 years (Martins, do Amaral, Tavares, Martins, Gonçalves, & Dias, 2016).

Additionally, certain voice disorders are more prevalent within treatment-seeking populations. A study conducted by Thijs et al. (2020) examined eight distinct diagnoses and found that mid-membranous lesions (MMLs) and nonspecific dysphonia (NSD) were the most commonly occurring diagnoses in a treatment-seeking population (Thijs, Knickerbocker, & Watts, 2020). Further, the prevalence of certain diagnoses is dependent on the gender and age of the patient. In a study of children ranging from one to 18 years, MMLs predominated with 59%

of the sample being diagnosed with vocal fold nodules, followed by vocal cysts. Acute laryngitis, formally classified as NSD, was the third most common disorder (Martins, do Amaral, Tavares, Martins, Gonçalves, & Dias, 2016). These findings are consistent with a previous study where vocal fold nodules were found to be the most prevalent in the age group zero to 14 years (Van Houtte, Van Lirder, D'Haeseleer, & Claeys, 2010). As for the adult population, gastroesophageal reflux disease (GERD) with laryngitis, polyps, and cysts are most prevalent in the age group 25 to 44 years, and edema or Reinke's edema are most prevalent in the age group of 25 to 64 years (Van Houtte, Van Lirder, D'Haeseleer, & Claeys, 2010).

1.4. Clinical Assessment of Voice

The diagnostic process to clinically assess voice consists of three primary areas: preliminary information or case history, instrumental analysis of voice function, and perceptual analysis of voice. Instrumental analysis of voice function can be divided into categories of visualization, acoustic measures, and aerodynamic measures. Perceptual analysis of voice can also be divided into categories of auditory-perceptual judgment and self-perception of the patient. All areas of clinical assessment of voice are described in detail in future sections.

Preliminary Information or Case History

Basic documentation of preliminary information and case history includes medical history and current health conditions of the patient. For voice in particular, it is crucial to examine and document allergies or allergy symptoms (specifically environmental or seasonal allergies), current or previous respiratory conditions, dyspnea with speech and/or exertion, reflux symptoms, dysphagia concerns or symptoms, and current medications.

In clinical assessment of voice, clinicians can begin to understand how the patient perceives their voice, what their complaint is, and what the characteristics of the voice problem are. By inquiring about and understanding a patient's medical history and current health conditions, clinicians can better predict what factors may be contributing to changes in voice, as well as any interventions or medications that are currently being trialed to treat those symptoms. Additionally, clinicians can perform subjective, auditory-perceptual measures to evaluate the patient's voice quality while simultaneously obtaining the patient's subjective interpretation of the voice problem. Gathering a subjective case history can also allow clinicians to probe whether or not changes have occurred since the patient was previously seen by the referring physician. This can also tell clinicians what actions the patient has taken to try and remedy the voice disorder, such as "home remedies," non-prescribed medications, interventions found online, or even previous professional intervention by SLPs.

Other important factors to note in clinical assessment are problem onset, consistency, and vocal demand. Dysphonia either presents with a gradual onset, generally through habituated vocal misuse or with a sudden onset as a result of neurological insult or iatrogenic surgical injury. Additionally, symptoms may vary between daily activities and situations of stress and are dependent on the time of day. Two questions often inquired in a clinical assessment of voice are whether patients perceive their voice as worse in the morning or afternoon and whether or not their symptoms improve with rest. It is also important to note the patient's current vocal demand and how that may be impacted by their occupation and/or familial situations (Watts & Awan, 2019).

Instrumental Analysis

Visualization

Laryngoscopy can be performed to visualize the vocal folds using a mirror, a rigid endoscope, or a flexible transnasal endoscope. Typically, rigid transoral endoscopy can provide the best magnification, illumination, and mucosal detail of the laryngeal cavity and vocal folds. Considered the “gold standard” for videostroboscopy, rigid transoral endoscopy is performed with a 70-degree telescope inserted into the oral cavity. In contrast to flexible transnasal laryngoscopy which allows patients to phonate on various pitches and vowels during the exam, rigid transoral endoscopy can only be performed on the vowel “ee” (/i/) due to contact with anatomical structures during the exam (Hapner & Johns, 2021).

A stroboscopic exam added onto laryngoscopy is the component that is most important for visualization of the vocal fold behavior during phonation. This utilization of a synchronized, flashing light can be passed through a flexible or rigid scope and allows for visualization of vocal fold vibration during sound production. Videostroboscopy is preferable as it allows the human eye to see several components of the vocal fold movement necessary for healthy phonation such as symmetry of movement, amplitude of lateral excursion, periodicity, mucosal wave motion, glottal closure, phase symmetry and closure, free edge contour, and adynamic or supra-glottic activity (Hapner & Johns, 2021).

Acoustic Measures

Contemporary acoustic measures of voice function are fundamental frequency, amplitude, cepstral peak prominence, multidimensional measures, and measures of frequency, amplitude, and spectral variability. Fundamental frequency is utilized frequently in clinical settings as it is a simple measure of the number of times the vocal folds complete a vibratory cycle within a second. Fundamental frequency is measured in hertz (Hz), is perceived as pitch, and is variable dependent on gender and age. High-tech options to determine acoustic parameters of voice function often involve the utilization of computer software. Software that can be used for acoustic voice analysis includes Praat, Dr. Speech, VoxMetria, DiVas, Lingwaves, and Multi-Dimensional Voice Program (MDVP), all of which were utilized in previous studies examining the relationship between objective and subjective measures of voice (Zhao, Nguyen, Salvador, & O'Rourke, 2020). This study in particular utilized the Acoustic Voice Quality Index (AVQI), a high-tech acoustic measure that will be further elaborated on in the methods section.

Aerodynamic Measures

As with acoustic measures of voice function, there are both high- and low-tech options for measures of aerodynamic qualities of voice. Low-tech options to measure aerodynamics involve the timing of phonation with a stopwatch. In contrast, high-tech options often involve the utilization of computerized software, such as the KayPentax Phonatory Aerodynamic System. Other aerodynamic measures of voice function include maximum phonation time (MPT), transglottal airflow rate, subglottal pressure, glottal resistance, and vital capacity. Because MPT was utilized within this study, it will be explained in further detail within the methods section.

Perceptual Analysis of Voice

Auditory Perceptual Judgement

Auditory perceptual judgment of the voice involves listening to and documenting perceptual characteristics of the patient's voice and using this information to compare characteristics of the voice of the patient to a "typical" voice, or one that is within normal limits (WNL). Auditory perceptual judgment is important in clinical assessment as this is often the reason that the patient has been referred, and results can be used as a primary gauge for therapy recommendations and patient motivation. Auditory perceptual measures are easily accessible to clinicians – at the most basic level, voice can be assigned to categories such as mild, moderate, or severe with a description of the perceptual features. A statement such as "the patient presents with mild dysphonia characterized by intermittent roughness, breathiness, and mild strain" represents an example of auditory perceptual judgment that can be performed should the clinician not have access to standardized auditory perceptual measures or equipment necessary to make a formal judgment on the severity of dysphonia.

Two perceptual data collection tools that are commonly utilized are the GRBAS scale (or the GRBASI scale, an extended edition) as well as the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Each letter of GRBAS represents a different domain of voice: G representing grade, R representing roughness, B representing breathiness, A representing asthenia or weakness, and S representing strain. With this method a score of zero to three is given for each letter with zero being normal, one being mild, two being moderate, and three being severe. An alternate tool to evaluate the voice perceptually is the CAPE-V, which was developed to help standardize clinical auditory-perceptual assessment of voice and to describe the severity of perceptual attributes in a manner that facilitates communication across clinicians.

Self-Perception of the Patient

The other category of perceptual analysis in clinical assessment of voice is self-perception of the patient, which is typically measured through Patient-Reported Outcome Measures (PROMs). PROMs allow clinicians to gauge the self-perception of the patient, and to understand the patient's perception of how the possible vocal dysfunction may be affecting their activities of daily living (ADLs) and overall quality of life. Results from these measures may also be indicative of the patient's self-awareness of the disorder, the possible level of stress and anxiety on the patient as a result, and the potential degree of patient motivation to follow through on treatment recommendations. PROMs are typically developed to determine one of three different measures: disability, the impact of performance due to impairment; handicap, the impact of impairment or disability on social, environmental, or economic functioning; or impairment-level, the measure of phonatory function such as listener ratings of dysphonia severity, as well as acoustic and aerodynamic measures.

Patient-Reported Outcome Measures are often administered in order to measure voice handicap and understand the impact of a voice disorder on an individual's everyday life. In a comprehensive systematic review of the literature on voice-related PROMs, over

30 different measures of voice handicap were analyzed for validity and reliability and revealed various categories of voice-related PROMs such as coping, quality of life, handicap, vocal performance, vocal impairment, vocal fatigue, voice quality, self-efficacy, and work productivity. Although there are many voice-related PROMs that are frequently utilized in clinical settings, the most popular among these are the Voice Handicap Index-30 (VHI-30) and Voice Handicap Index-10 (VHI-10). The VHI-30 is the most studied and most popular patient-based instrument and is a psychometrically validated tool developed for the measurement of the psychosocial effects of voice disorders (Watts & Awan, 2019). The VHI-30 was initially released in 1997 with the intent to develop a psychometrically robust voice disability/handicap inventory that could be used with patients exhibiting a variety of voice disorders. In 2004, the VHI-30 was redesigned with the intent to develop a shortened version, resulting in the VHI-10. The VHI-30 and VHI-10 are discussed in further detail within the methods section.

CHAPTER II: REVIEW OF LITERATURE

2.1. Relationship Between Perceptual and Objective Measures

Historically, studies examining the relationship between perceptual and objective measures of voice were often unable to determine a correlation between the two. For example, Cheng and Woo (2010) found that acoustic measures of fundamental frequency, jitter, shimmer, and noise-to-harmonic ratio (NHR), as well as the aerodynamic measure of maximum phonation time (MPT), did not significantly correlate with VHI scores. Similarly, Wheeler et al. (2006) and Hsiung et al. (2002) also found that individual components of the VHI did not consistently or significantly correlate with acoustic measures of voice such as fundamental frequency, jitter, and shimmer (Wheeler, Collins, & Sapienza, 2006; Hsiung, Pai, & Wang, 2002). While the majority of prior studies utilized only sustained vowel phonation, Wheeler et al. (2006) utilized both sustained vowel phonation and connected speech. Interestingly, fundamental frequency was the only acoustic measure to significantly correlate with overall VHI score within the connected speech analysis (Wheeler, Collins, & Sapienza, 2006).

Recent studies have found significant correlations, however, between perceptual measures and objective measures of voice. For example, both Karlsen et al. (2020) and Lopes et al. (2017) have found a weak to moderate relationship between clinical measures of voice and patient self-perception (Karlsen, Sandvik, Heimdal, & Aarstad, 2020; Lopes, da Silva, Simoes, Evangelista, Silva, Almeida, & de Lima-Silva, 2017). It should be noted that all of the aforementioned studies utilized the same aerodynamic measure: maximum phonation time (Karlsen, Sandvik, Heimdal, & Aarstad, 2020; Lopes, da Silva, Simoes, Evangelista, Silva, Almeida, & de Lima-Silva, 2017; Cheng & Woo, 2010; Wheeler, Collins, & Sapienza, 2006; Hsiung, Pai, & Wang, 2002); for acoustic measures, these studies used jitter, shimmer, harmonic-to-noise ratio (HNR), noise-to-harmonic ratio (NHR), and fundamental frequency.

A recent meta-analysis of the association between VHI-30/VHI-10 and objective voice parameters found that one limitation in determining the relationship between perceptual and objective measures is that prior studies evaluating the relationships between patient-reported and acoustic or aerodynamic voice analysis outcomes are scarce with variable results (Zhao, Nguyen, Salvador, & O'Rourke, 2020). Houle & Johnson (2021) supplement the limitations put forth in the meta-analysis performed by Zhao et al. (2020) as they suggest that variance is not accounted for in clinical measures of voice – that is, the impact of unique environmental and personal factors on a patient's self-perceived severity and participation in daily life (Houle & Johnson, 2021). Further, within their study, they found that, for both males and females, auditory perceptual measures of voice quality and functioning contributed significantly to patient-reported outcome measures of voice (Houle & Johnson, 2021).

A limitation of previous studies, however, is that acoustic and aerodynamic measurements were extracted primarily from sustained vowel phonation, as opposed to both sustained vowel phonation and connected speech. In voice analyses, a connected speech sample is preferable as the quantity and quality of voiced fragments increases. Specifically, a connected speech sample differs from sustained vowel phonation in that connected speech is a continuous stream of speech which alters how words or syllables are produced, the way phrases run together, and the emphasis placed on each word. Sustained vowel phonation is equally important in voice analysis, but it cannot provide the same contextual information as that of connected speech.

This study differs from previous studies as, within this study, both connected speech and sustained vowel phonation will be utilized. In particular, both the CAPE-V and AVQI utilize connected speech and sustained vowel phonation, and both will be utilized within this study. The aerodynamic measurement utilized within this study, MPT, measures only sustained vowel phonation as the task is only for the patient to hold a vowel, "ah", for as long as he or she can until they no longer have the breath to support phonation. The VHI-30 and VHI-10 are patient-reported measures to indicate overall impact of dysphonia on the patient's everyday life and therefore do not consider the type of speech sample collected. This study differs from previous studies, particularly in the utilization of AVQI for the acoustic measure, which includes various acoustic parameters of voice both in connected speech and sustained vowel phonation. While previous studies have utilized three to four different acoustic parameters of voice, the AVQI utilizes six different acoustic parameters of voice; the details of such will be elaborated on in the methods section (Karlsen, Sandvik, Heimdal, & Aarstad, 2020; Lopes, da Silva, Simoes, Evangelista, Silva, Almeida, & de Lima-Silva, 2017; Cheng & Woo, 2010; Wheeler, Collins, & Sapienza, 2006; Hsiung, Pai, & Wang, 2002).

2.2. Purpose of the Study

The importance of further understanding the relationship between the VHI-30/VHI-10 and objective voice measurements is supported by current clinical practice, which considers both objective measurements and quality of life when selecting treatment for patients. By increasing our understanding of the personal impact of dysphonia as compared to acoustic and aerodynamic measurements of voice, we can prioritize interventions to increase the likelihood of improving objective parameters with the greatest influence on the patient's quality of life, as reported through subjective measurements such as the VHI-30/VHI-10. This topic and its clinical application, however, are limited by the amount and quality of current published data. Current evidence indicates that objective voice parameters are poorly correlated with the VHI-30/VHI-10; however, there are limitations in the data collected by prior studies. Our understanding of the correlation between objective measurements and patient-perceived voice quality may be improved by utilizing a different acoustic measure of voice than what has been historically used. The understanding of the relationship between patient-perceived voice handicap and clinical measures of voice can also be improved by considering both connected speech and sustained vowel phonation in the analysis of correlation.

Regardless, further research is needed to determine if changes in acoustic and aerodynamic parameters of voice are associated with VHI-30/VHI-10 scores and an overall change in quality of life for the patient. The purpose of this study is to further examine the correlation between acoustic measures of voice and the VHI-30/VHI-10, specifically utilizing the AVQI which encompasses various parameters of voice and considers both sustained vowel phonation and connected speech in its analysis. In aiming to further understand the relationships between a patient's self-perceived voice handicap and clinical measures of voice, current knowledge can be extended to improve the clinical ability to select treatment methods accordingly, further consider the impact of voice disorders on a patient's daily life, and to improve the understanding and prediction of objective outcomes in the initial evaluation.

2.3. Research Questions

This study addressed the following research questions: (1) What is the effect of disorder type on the measurement of Maximum Phonation Time (MPT), Acoustic Voice Quality Index (AVQI), Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), and Voice Handicap Index scores (VHI – VHI-30 version and VHI-10 version); and (2) what is the strength of the relationship between pre-treatment measures of the Acoustic Voice Quality Index (AVQI) and the Voice Handicap Index-30 (VHI-30)/Voice Handicap Index-10 (VHI-10); and (3) what is the strength of the relationship between the pre-treatment measures of Maximum Phonation Time (MPT) and the VHI-30/VHI-10; and (4) what is the strength of the relationship between the pre-treatment perceptual measure of Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) and the VHI-30/VHI-10; and (5) what is the relationship of measures AVQI, MPT, CAPE-V, and VHI-30/VHI-10 among different types of voice disorders?

2.4. Hypotheses

For research question one we hypothesized that disorder type would have an effect on the measurements of MPT, AVQI, CAPE-V, and VHI-30/VHI-10 scores, and that effect would be strong and significant for both the VHI-30 and VHI-10 subgroup. For research questions two and three, we hypothesized that correlations between the pre-treatment measures of AVQI compared to VHI-30/VHI-10, and the pre-treatment measures of MPT compared to VHI-30/VHI-10 would be significant and strong. For research question four, we hypothesized that correlations between the pre-treatment measure of CAPE-V compared to VHI-30/VHI-10 would not yield a significant relationship. We hypothesized that for research question five, there would be a relationship between disorder and pre-treatment measures of AVQI, MPT, CAPE-V, and VHI-30/VHI-10, and that relationship would be significant and strong.

CHAPTER III: METHODOLOGY

3.1. Participants

The methodology for this study was approved by the Texas Christian University Institutional Review Board. The study utilized data from an existing database from a private practice community voice clinic. A previous analysis of the database was reported by Thijs, Knickerbocker, and Watts (2020), as well as Watts and Knickerbocker (2019). All patient referrals to the voice clinic were from either (1) community otolaryngology practices, (2) community SLPs, or (3) self-referrals. Data extracted for this investigation represented information obtained during a specialty voice evaluation by the SLP, along with records from the patients' otolaryngologist. Previous or subsequent otolaryngology examinations were completed for each patient prior to the development of any voice treatment plan. Medical diagnoses were established via the otolaryngology and specialty voice clinical evaluations. Laryngeal imaging via videostroboscopy was completed by either a referring otolaryngologist or the SLP. The International Classification of Diseases, 10th Revision, was utilized for guidance on diagnosis codes, consistent with previous studies.

Patient data were collected from charts associated with 805 patients receiving a voice evaluation across 50 consecutive months, regardless of age and disorder type. From the original data set, 166 cases were removed for the following reasons: (a) if an alternative assessment was given to determine voice handicap, (b) if only the VHI score was reported, or (c) if no VHI score was reported. Because part of the sample was administered the VHI-30 before the practice switched to administering the VHI-10, data was split into VHI-30 and VHI-10 subgroups for all data analyses. Sample size and visualization of disorder distribution for each subgroup will be outlined in Section 4.1.

3.2. Measures

The independent variables within this study were disorder type, as well as age and sex outlined within descriptive data analysis. The measures or dependent variables utilized for this study were (1) Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), (2) maximum phonation time (MPT), (3) Acoustic Voice Quality Index (AVQI), and (4) the Voice Handicap Index-30 (VHI-30) or Voice Handicap Index-10 (VHI-10). VHI subgroup was also incorporated as a factor within the study to separately assess how other measures associated with these two forms of the VHI instrument.

The CAPE-V is a subjective perceptual measurement. In particular, the CAPE-V fulfills the role of auditory perceptual judgment performed by the clinician. CAPE-V examines perceptual domains associated with overall severity, roughness, breathiness, strain, pitch, and loudness. The CAPE-V also allows for the measurement of other characteristics that may occur, such as diplophonia, asthenia, or tremor. Each domain on the CAPE-V is measured on a 100-millimeter line. With this, the clinician indicates the degree of perceived deviance from normal for each parameter with a tick-mark, and regions of mildly deviant (MI), moderately deviant (MO), and severely deviant (SE) are indicated to aid in the placement of the tick-mark. A rating greater than zero indicates perceptions that might fall outside the range of normal for that perceptual domain.

The MPT is considered an aerodynamic measurement. MPT is the longest time that a patient can sustain a vowel sound after a maximum inhalation. Typically, a decreased MPT is observed in dysphonia patients compared to healthy controls, with 20 seconds of phonation being within normal limits for adults (Zhao, Nguyen, Salvador, & O'Rourke, 2020). To date, MPT is the most commonly used aerodynamic measurement.

The AVQI is an acoustic analysis that is closely associated with perceived voice quality. AVQI is a multiparametric tool used to objectively quantify voice quality and severity. It is unique to acoustic analysis of voice in that it was designed to permit objective assessment of dysphonia on sustained vowels and connected speech whereas previous acoustic techniques were limited to sustained vowel phonation. AVQI is calculated by a weighted combination of six different acoustic markers to yield a single number that correlates with overall dysphonia severity. The output score for AVQI ranges from zero to 10, where more severely perceived dysphonia corresponds with a higher score. Within the English language, a number equal to or greater than 3.46 is associated with perceptual impressions of dysphonia (Maryn, De Bodt, & Roy, 2010). It should be noted that the reported value is normed in Dutch, however, this value is utilized for English as there is no normative study for AVQI within the English language.

The VHI-30 and VHI-10 are two forms of a similar instrument used to measure patients' self-perceptions of voice handicap. The VHI-30 was developed with the intent to quantify the patients' perception of disability due to voice impairment. Consisting of 30 questions, the resulting score of the VHI-30 is a range from zero to 120 with the latter representing the maximum perceived disability due to voice difficulties. The VHI-30 was designed to take less than five minutes to complete, without assistance from a clinician, and is intended to be sensitive to the patient's perception of change in severity following one or more interventions. While not formally analyzed within this study, demographic factors such as age, occupation, voice demands, and duration of the voice problem have been shown to affect the VHI-30 score, ultimately indicating its sensitivity to variability across patient demographics (Rosen, Murry, Zinn, Zullo, & Sonbolian, 2000).

In 2004, the VHI-30 was adapted into a shorter version, the VHI-10, with the intent to determine patient-perceived self-handicap at both initial evaluation and throughout longitudinal assessment of patients with voice disorders. (Jacobson, Johnson, Grywalski, Silbergleit, Jacobson, Benninger, & Newman, 1997; Rosen, Lee, Osborne, Zullo, & Murry, 2004). The VHI-10 consists of 10 questions with the resulting score ranging from zero to 40, with the latter representing the maximum perceived disability due to voice difficulties. To date, the VHI-30 and VHI-10 represent primary tools used to measure treatment outcomes in patients receiving voice therapy, with the majority of clinics using either the VHI-30 or the VHI-10 as a measure for both initial evaluation and discharge (Thijs, Knickerbocker, & Watts, 2020).

As an ad-hoc comparison, the disorder types of patients were stratified into different categories, using a similar approach as previously published studies (Watts & Knickerbocker, 2019; Thijs, Knickerbocker, & Watts, 2020). This resulted in eleven distinct groups of disorders:

(1) vocal fold insufficiency [paresis or paralysis – VFI], (2) atrophy or bowing, (3) neurological [tremor, laryngeal spasm, or dyskinesia], (4) other [patients presenting only with edema, reflux, or globus sensation], (5) multifactorial [concomitant dysphagia, GERD, dysarthria, or multiple diagnoses of voice disorders], (6) patients who are transgender, (7) non-specific dysphonia [dysphonia unspecified – NSD], (8) vocal cord dysfunction [VCD], (9) mid-membranous lesions [MMLs], (10) muscle tension dysphonia [MTD], and (11) patients with a neoplasm [pre-cancerous diagnoses or benign/malignant tumors].

3.3. Data Collection

To address the research questions, the following data were extracted from the larger sample of patients: (1) auditory perceptual ratings of voice quality (CAPE-V), (2) maximum phonation time (MPT), (3) Acoustic Voice Quality Index (AVQI) scores, and (4) Voice Handicap Index scores (VHI-30 or VHI-10). The auditory perceptual rating of voice quality was performed by the SLP. All data represent measurements at pre-treatment (baseline) periods, prior to the application of any voice therapy.

3.4. Statistical Analysis

All data analysis for the sample was performed using SPSS (version 27.0, Armonk, NY: IBM Corp). Prior to statistical analysis, four data points were removed from the data set due to apparent coding errors (two from AVQI data, two from VHI-10 data). Following this, 10 outliers were removed from MPT, and six outliers were removed from AVQI data sets. No outliers were removed from CAPE-V data. Based on the SPSS definitions, outliers were defined as any data value that was outside of the third quartile or the first quartile at 1.5 times the interquartile range.

Extreme outliers were considered to be any data value that was outside of the third quartile or first quartile, at three times the interquartile range. Because VHI-30 and VHI-10 utilize different scale ranges the sample was divided into two subgroups: patients who had completed the VHI-30 and patients who had completed the VHI-10. In order to address multiple testing and enhance the reliability and reproducibility of results, the default significance level alpha = 0.05 was adjusted to alpha = 0.025. Additionally, one VCD case within the VHI-30 group was coded as "other" due to low group membership for Post Hoc analysis.

For Research Question 1 eight separate ANOVAs were applied to the data to determine the effect of disorder on each measurement (MPT, AVQI, CAPE-V, VHI-30/VHI-10 scores). Four ANOVAs were applied to the VHI-30 subgroup, and four ANOVAs were applied to the VHI-10 subgroup. Bivariate correlation was performed in order to determine the correlation between pre-treatment measures of AVQI and VHI-30/VHI-10 scores (Research Question 2), MPT, and VHI-30/VHI-10 scores (Research Question 3), and CAPE-V and VHI-30/VHI-10 scores (Research Question 4). Bivariate correlation was also performed for Research Question 5 in order to determine the correlation between disorder type and measures of AVQI, MPT, CAPE-V, and VHI scores (VHI-30/VHI-10). As with previous analyses, data was split into VHI-30 and VHI-10 subgroups prior to analysis. Due to the small occurrence, two cases of MTD were coded as "other" within the VHI-10 subgroup.

CHAPTER IV: RESULTS

4.1. Descriptive Data

While not formally addressed within the research questions of this study, descriptive data was analyzed in order to better understand the sample demographics and overall scoring across assessments within subgroups. Descriptive data of age and sex (assigned at birth) is outlined in Table 1. Within the sample, the oldest average age was found in the atrophy/bowing disorder group at 73.77 (SD 9.46) and the youngest average age was in the transgender group at 28.37 (SD 11.20). There were predominantly more females in the study, making up 68.3% of the sample. The largest disorder group was non-specific dysphonia, making up 38.6% of the sample.

Table 1[1]

Age and sex (assigned at birth) distribution of the sample across disorder groups

		Age	Sex[2]	
	n	Mean (SD)	Male	Female
VFI	84 (13.6%)	58.86 (16.60)	16 (19.0%)	68 (81.0%)
Atr/Bow	31 (5.0%)	73.77 (9.46)	16 (51.6%)	15 (48.4%)
Neuro	12 (1.9%)	63.75 (7.36)	3 (25.0%)	9 (75.0%)
Multi	54 (8.7%)	59.39 (17.03)	17 (31.5%)	37 (68.5%)
Trans	30 (4.9%)	28.37 (11.20)	30 (100.0%)	0 (0.0%)
NSD	239 (38.6%)	55.21 (18.34)	73 (30.5%)	166 (69.5%)
VCD	14 (2.3%)	52.79 (19.55)	2 (14.3%)	12 (85.7%)
MMLs	100 (16.2%)	46.12 (15.95)	23 (23.0%)	77 (77.0%)

[1] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm

[2] Sex assigned at birth

NSD	17 (2.8%)	50.65 (16.19)	2 (11.8%)	15 (88.2%)
Neo	13 (2.1%)	57.46 (11.43)	3 (23.1%)	10 (76.9%)
Other	25 (4.0%)	55.83 (17.58)	11 (44.0%)	14 (56.0%)
Total	619 (100.0%)	54.21 (18.56)	196 (31.7%)	423 (68.3%)

Descriptive data of sample size and disorder occurrence for the entire sample, as well as VHI-30 and VHI-10 subgroups, are outlined in Table 2. Visual representations of disorder occurrence for the entire sample, the VHI-30 subgroup, and the VHI-10 subgroup are outlined in Graph 1, Graph 2, and Graph 3 respectively.

Table 2[3]

Sample size and disorder occurrence for sample and VHI-30/VHI-10 subgroups

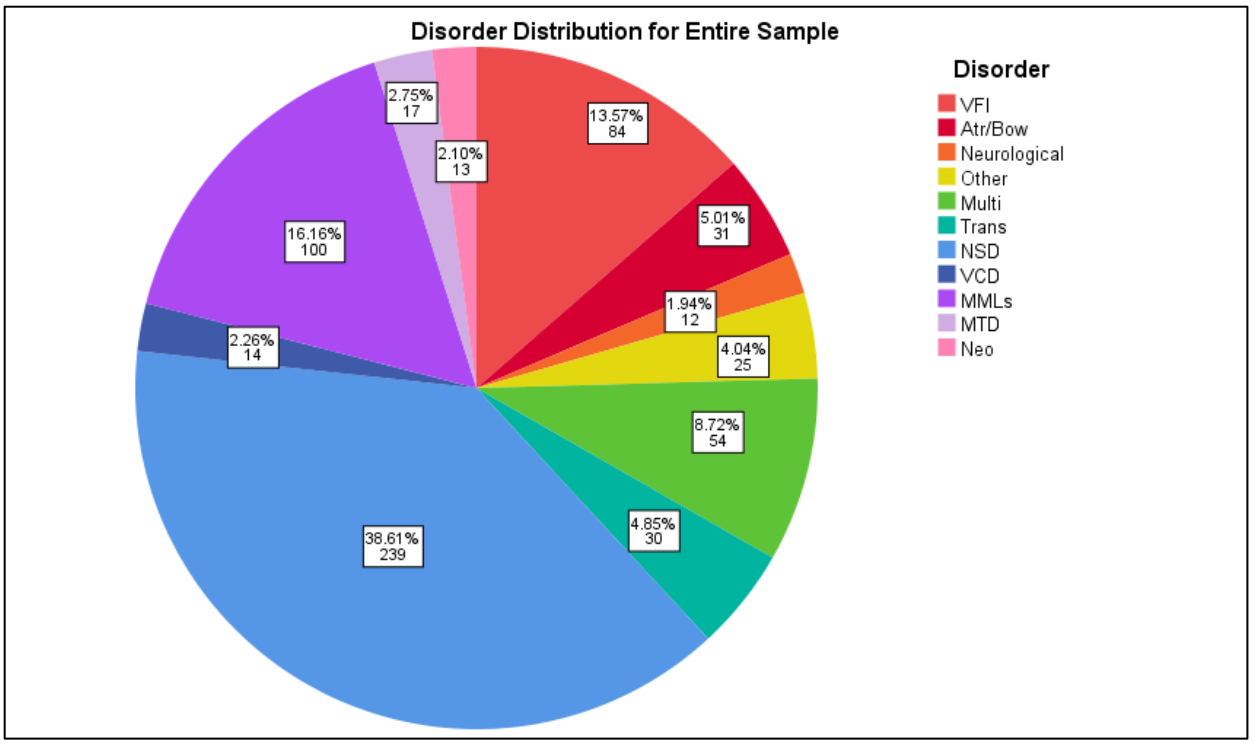
	VHI-30		VHI-10		Sample	
	n	Percentage	n	Percentage	n	Percentage
VFI	33	15.6%	51	12.5%	84	13.57%
Atr/Bow	27	12.7%	4	1.0%	31	5.01%
Neuro	3	1.4%	9	2.2%	12	1.94%
Other	8	3.8%	17	4.2%	25	4.04%
Multi	21	9.9%	33	8.1%	54	8.72%
Trans	8	3.8%	22	5.4%	30	4.85%
NSD	47	22.2%	192	47.2%	239	38.61%
VCD	3	1.4%	11	2.7%	14	2.26%

[3] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm

MMLs	42	19.8%	58	14.3%	100	16.16%
MTD	15	7.1%	2	0.5%	17	2.75%
Neo	5	2.4%	8	2.0%	13	2.10%
Total	212	100%	407	100%	619	100%

Graph 1[4]

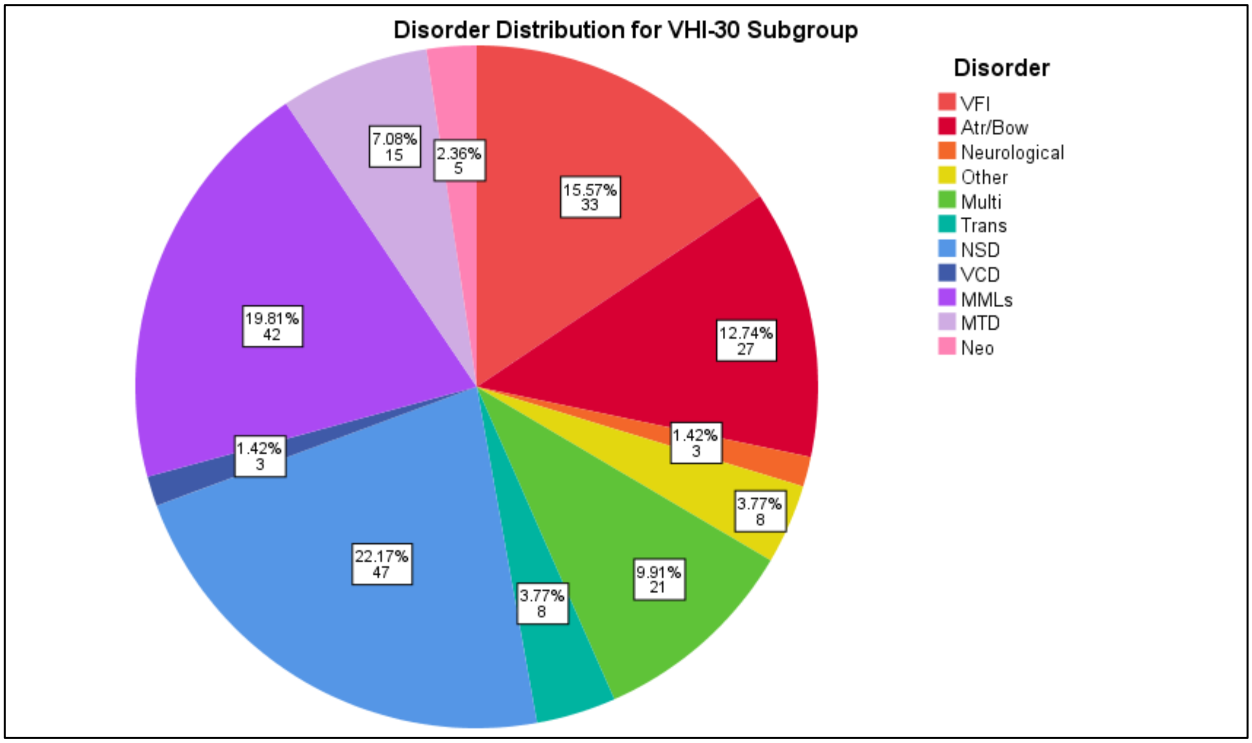
Visual distribution of disorder type for entire sample



[4] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm

Graph 2[5]

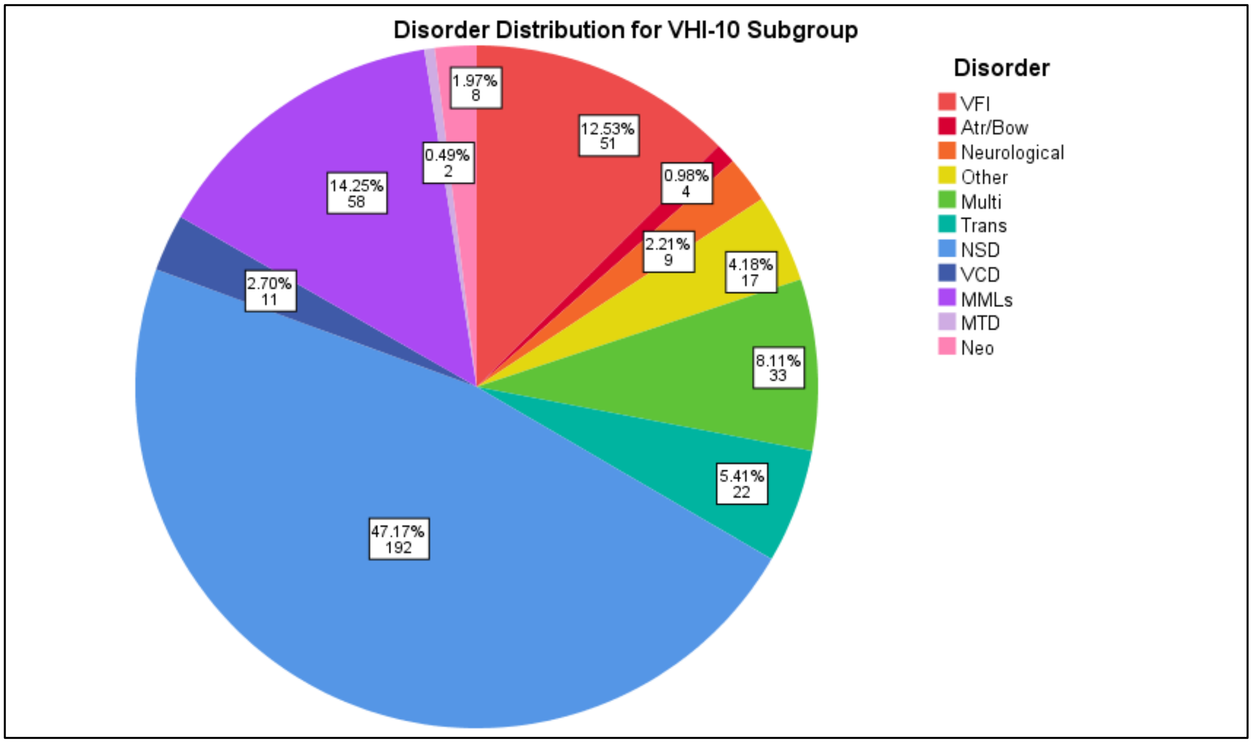
Visual distribution of disorder type within VHI-30 subgroup



[5] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, VHI-30 = Voice Handicap Index-30

Graph 3[6]

Visual distribution of disorder type within VHI-10 subgroup



In order to understand the similarities and differences of assessment scores for each subgroup, the descriptive data for VHI-30 and VHI-10 compared to MPT, AVQI, and CAPE-V can be found in Table 3. In general, participants who completed the VHI-10 had a higher average score on the AVQI (4.09, SD 1.81) than those who completed the VHI-30 (3.94, SD 1.85). VHI-10 participants also had a higher average score for the CAPE-V (51.36, SD 27.08) compared to VHI-30 participants (44.45, SD 27.80). On the contrary, participants who completed the VHI-30 had an increased MPT (18.56, SD 10.41) as compared to those who completed the VHI-10 (16.69, SD 9.12).

[6] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, VHI-10 = Voice Handicap Index-10

Table 3[7]

Mean and Standard Deviation for VHI-30 and VHI-10 across assessments

	VHI-30		VHI-10	
	n	Mean (SD)	n	Mean (SD)
MPT	11	18.56 (10.41)	202	16.69 (9.12)
AVQI	184	3.94 (1.85)	393	4.09 (1.81)
CAPE-V	209	44.45 (27.80)	196	51.36 (27.08)
VHI-30	212	49.90 (27.29)	-	-
VHI-10	-	-	405	18.53 (10.53)

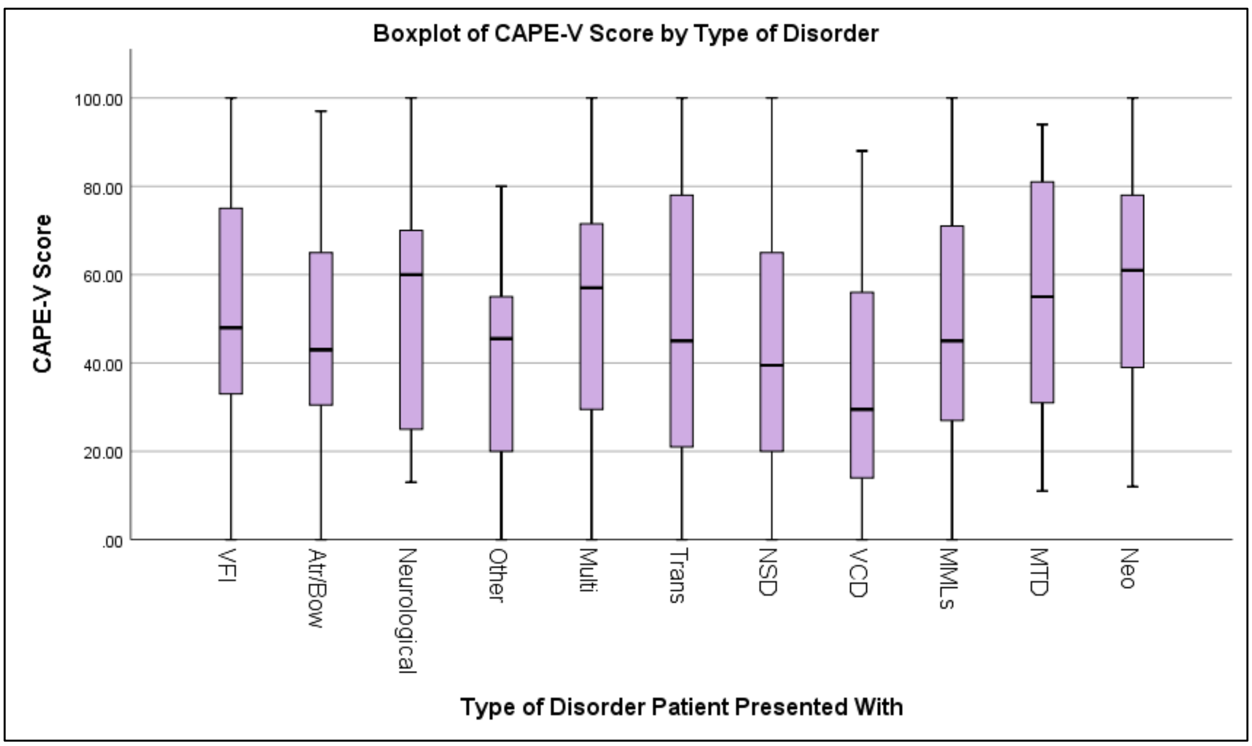
Visual representations of descriptive data for each assessment, by disorder group, can be found in Graph 4 (CAPE-V), Graph 5 (MPT), and Graph 6 (AVQI). It should be noted that within these graphs data is collapsed to include VHI-30 and VHI-10 patients. Within CAPE-V scores (Graph 4) the disorder type with highest median was neoplasm (61.00, SD 29.17), followed by the neurological disorder group (60.00, SD 30.57); the lowest median within CAPE-V scores was found in the VCD disorder group (29.50, SD 30.30). The highest averages among CAPE-V scores were found in the neoplasm (59.56, SD 29.12) and neurological (56.22, SD 30.57) disorder groups, and the lowest averages were found in the “other” (41.57, SD 23.76) and NSD (42.17, SD 27.10) disorder groups. For MPT scores, (Graph 5) the neoplasm disorder group (21.00, SD 10.52) and the transgender disorder group (21.00, SD 9.81) had the highest medians. The lowest median for MPT scores was found within the VFI disorder group (10.00, SD 10.53),

[7] Note: MPT = Maximum Phonation Time, AVQI = Acoustic Voice Quality Index, CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

aligning with what is known about the physiology of this disorder group. The VFI disorder group also had the lowest average among MPT scores (14.04, SD 10.53), whereas the transgender (23.89, SD 9.82) and neoplasm (20.29, SD 10.52) disorder groups had the highest averages. The VCD disorder group (5.00, SD 2.01) had the highest median across AVQI scores (Graph 6) whereas the "other" disorder group (3.09, SD 1.91) and the transgender disorder group (3.07, SD 1.43) had the lowest medians. Among AVQI scores, the VFI (4.53, SD 2.05) and atrophy/bowing (4.56, SD 1.58) disorder groups had the highest average; the lowest average was found in the transgender disorder group (3.06, SD 1.43).

Graph 4[8,9]

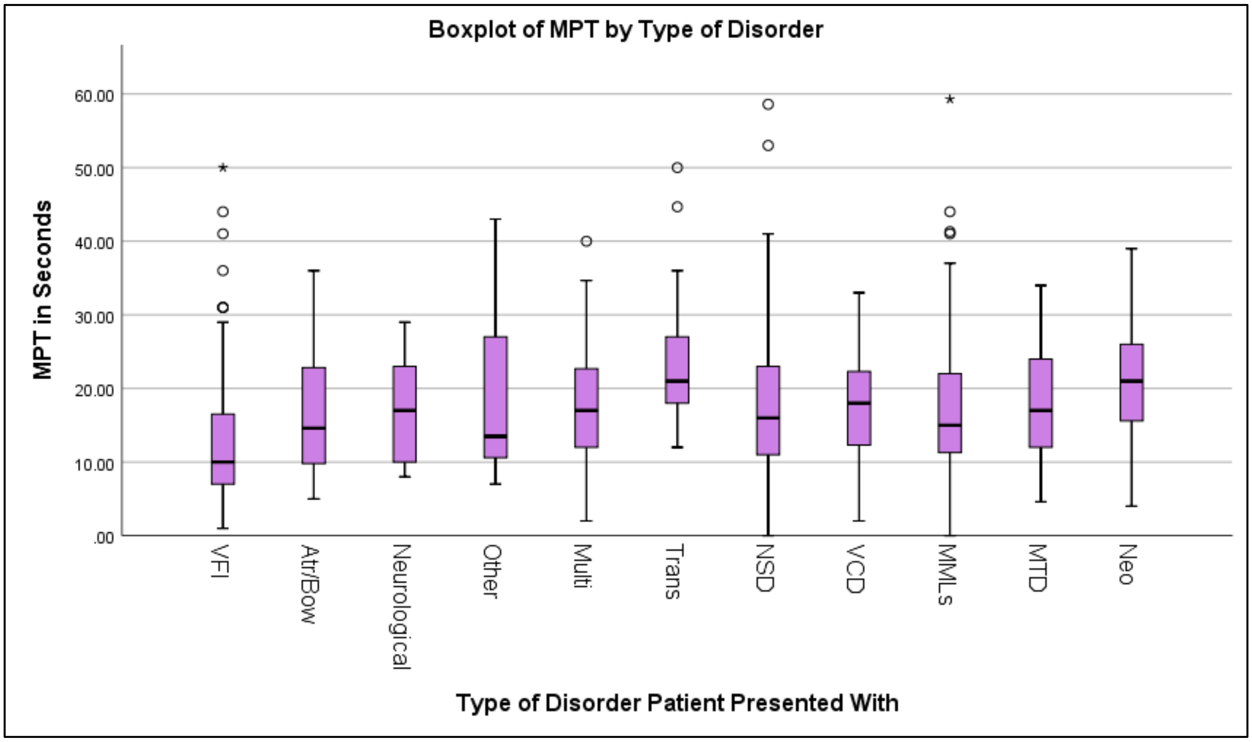
Visual distribution of CAPE-V scores by disorder type



[8] The data displayed contains both the VHI-30 and VHI-10 subgroups.
[9] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice

Graph 5[10, 11,12]

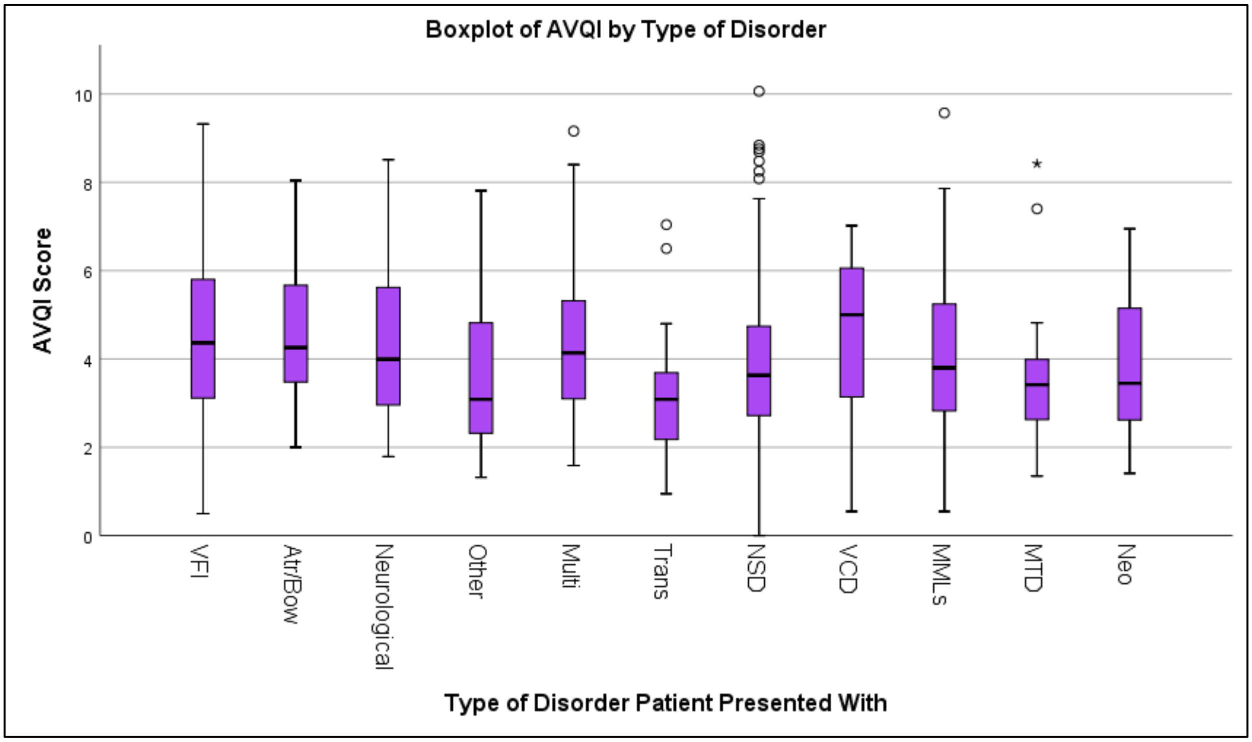
Visual distribution of MPT scores by disorder type



[10] The data displayed contains both the VHI-30 and VHI-10 subgroups.
[11] Within the graph, the colored area is considered to be the distance between the first quartile and the third quartile, otherwise known as the interquartile range (IQR). Outliers are represented as ○, indicating that the data value is 1.5 times the IQR range larger than the third quartile, or 1.5 times the IQR smaller than the first quartile. Extreme outliers are represented as *, indicating that the data value is 3 times the IQR range larger than the third quartile, or 3 times the IQR smaller than the first quartile.
[12] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, MPT = Maximum Phonation Time

Graph 6[13, 14, 15]

Visual distribution of AVQI scores by disorder type



[13] The data displayed contains both the VHI-30 and VHI-10 subgroups.
[14] Within the graph, the colored area is considered to be the distance between the first quartile and the third quartile, otherwise known as the interquartile range (IQR). Outliers are represented as ○, indicating that the data value is 1.5 times the IQR range larger than the third quartile, or 1.5 times the IQR smaller than the first quartile. Extreme outliers are represented as *, indicating that the data value is 3 times the IQR range larger than the third quartile, or 3 times the IQR smaller than the first quartile.
[15] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, AVQI = Acoustic Voice Quality Index

Descriptive data for each assessment by disorder group can be found in Table 4 (VHI-30) and Table 5 (VHI-10). For the VHI-30 subgroup (Table 4), the MPT average was highest in the transgender disorder group (28.10, SD 13.59) whereas the lowest average was in the VFI disorder group (12.99, SD 10.34). The average highest CAPE-V was in the MTD disorder group with 56.73 (SD 28.70), and the average lowest was in the transgender disorder group with 17.17 (SD 15.97). For the AVQI, the average highest score was within the VCD disorder group (5.35, SD N/A); however, with consideration that the VCD sample size is one, other disorder groups with a higher AVQI average were the multifactorial disorder group (4.74, SD 2.03) and the atrophy/bowing disorder group (4.63, SD 1.57). The lowest AVQI on average was the transgender group (2.42, SD 1.26). For VHI-30 scores the VFI disorder group had the highest average at 66.27 (SD 28.84) and, similar to AVQI, the lowest average was the VCD disorder group at 29.00 (SD 21.07). With consideration that the VCD sample size is one, other groups with a lower average VHI-30 score were the "other" disorder group (39.75, SD 26.59) and the atrophy/bowing disorder group (40.30, SD 18.15). A visual representation of descriptive data for VHI-30 scores can be found in Graph 7.

Table 4[16]

Mean and Standard Deviation of each assessment by disorder group for VHI-30

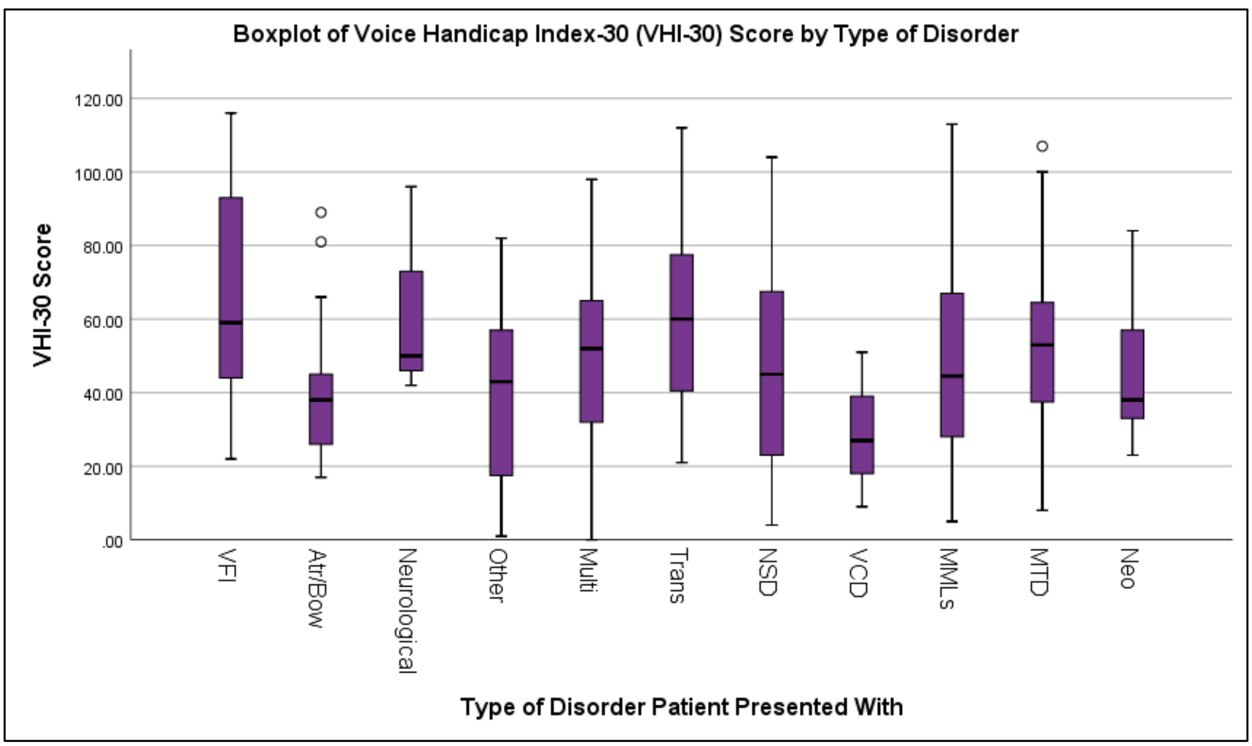
	MPT	CAPE-V	AVQI	VHI-30
VFI (n=30)	12.99 (10.34)	55.47 (28.85)	4.57 (1.97)	66.27 (28.84)
Atr/Bow (n=26)	16.75 (8.57)	45.74 (24.57)	4.63 (1.57)	40.30 (18.15)
Neuro (n=3)	13.67 (4.16)	65.00 (37.75)	3.89 (2.27)	62.67 (29.14)

[16] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, VHI-30 = Voice Handicap Index-30

Multi (n=18)	18.92 (7.38)	50.14 (30.11)	4.74 (2.03)	49.00 (27.28)
Trans (n=6)	28.10 (13.59)	17.17 (15.97)	2.42 (1.26)	61.13 (28.54)
NSD (n=39)	21.07 (11.60)	34.32 (24.37)	3.08 (1.40)	45.21 (27.50)
VCD (n=1)	17.20 (5.35)	26.67 (21.08)	5.35 (N/A)	29.00 (21.07)
MMLs (n=36)	19.16 (10.49)	42.98 (27.49)	3.74 (1.81)	48.60 (27.17)
MTD (n=12)	17.37 (7.68)	56.73 (28.70)	3.88 (2.10)	52.73 (27.09)
Neo (n=5)	25.12 (8.69)	47.20 (25.92)	3.76 (1.35)	47.00 (24.09)
Other (n=5)	20.65 (11.32)	43.00 (26.86)	4.33 (2.45)	39.75 (26.59)

Graph 7[17]

Visual distribution of VHI-30 scores by disorder type



[17] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, VHI-30 = Voice Handicap Index-30

For the VHI-10 subgroup (Table 5), the longest MPT was the transgender disorder group with 21.93 (SD 7.24) and the neoplasm disorder group had the average shortest MPT with 14.25 (SD 10.34). The disorder group with the average highest CAPE-V score was neoplasm (75.00, SD 28.30), and the average lowest CAPE-V score was the MTD disorder group (37.50, SD 23.33). For AVQI scores, the average highest scores were in the neurological disorder group with 4.58 (SD 1.87) and the VFI disorder group with 4.50 (SD 2.11). The average lowest AVQI scores were in the MTD disorder group and the transgender disorder group with 3.28 (SD 0.59) and 3.34 (SD 1.45), respectively. In regard to VHI-10 scores, the group with the average highest VHI-10 score was the neurological disorder group (23.11, SD 9.75). The groups with the average lowest VHI-10 scores were the transgender disorder group (12.77, SD 6.79) and the MTD disorder group (5.50, SD 0.71). A visual representation of descriptive data for VHI-10 scores can be found in Graph 8.

Table 5[18]

Mean and Standard Deviation of each assessment by disorder group for VHI-10

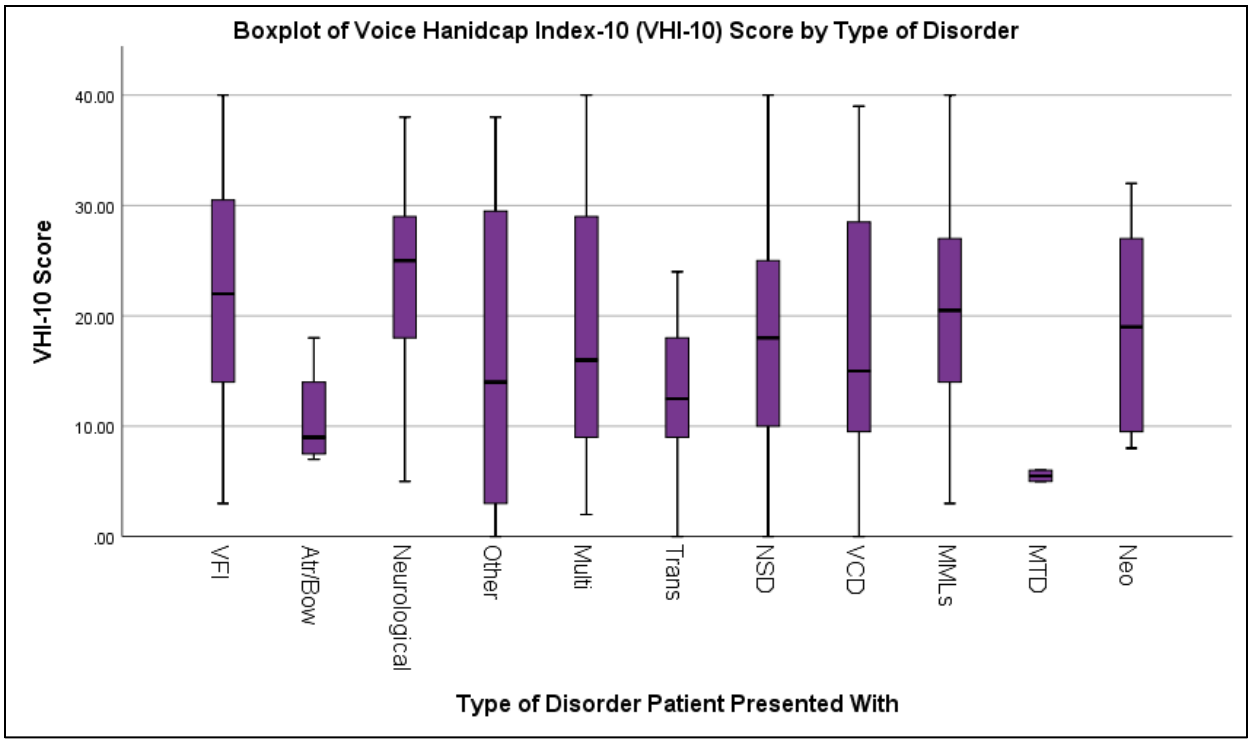
	MPT	CAPE-V	AVQI	VHI-10
VFI (n=29)	15.16 (10.77)	51.84 (24.98)	4.50 (2.11)	21.84 (9.84)
Atr/Bow (n=4)	16.25 (5.74)	58.50 (29.73)	4.07 (1.79)	10.75 (4.99)
Neuro (n=6)	19.17 (8.28)	51.83 (29.25)	4.58 (1.87)	23.11 (9.75)
Multi (n=14)	16.00 (9.50)	57.40 (26.34)	4.15 (1.75)	18.55 (11.82)
Trans (n=13)	21.93 (7.24)	59.87 (32.67)	3.34 (1.45)	12.77 (6.79)
NSD (n=69)	16.68 (8.31)	47.09 (27.71)	4.04 (1.80)	18.05 (10.51)

[18] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, VHI-10 = Voice Handicap Index-10

VCD (n=7)	18.43 (9.73)	42.00 (33.86)	4.36 (2.08)	18.18 (12.77)
MMLs (n=28)	15.69 (9.18)	55.39 (23.14)	4.23 (1.57)	20.43 (9.60)
MTD (n=2)	17.50 (16.26)	37.50 (23.33)	3.28 (0.59)	5.50 (0.71)
Neo (n=4)	14.25 (10.34)	75.00 (28.30)	3.88 (2.23)	18.88 (9.64)
Other (n=6)	16.00 (13.86)	39.67 (21.21)	3.48 (1.76)	15.81 (13.46)

Graph 8[19]

Visual distribution of VHI-10 scores by disorder type



[19] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, VHI-10 = Voice Handicap Index-10

4.2. Research Question 1: ***What is the effect of disorder type on measurement of Maximum Phonation Time (MPT), Acoustic Voice Quality Index (AVQI), Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), and Voice Handicap Index scores (VHI – VHI-30 version and VHI-10 version)?***

Eight separate ANOVAs were applied to the data to determine the effect of disorder on each measurement (MPT, AVQI, CAPE-V, VHI-30/VHI-10 scores). Four ANOVAs were applied to the VHI-30 subgroup and four ANOVAs were applied to the VHI-10 subgroup. Results are as follows: for the VHI-30 subgroup, a significant effect was found for MPT ($F[10]=2.37$, $p=0.011$, $\eta_p^2=0.11$, Observed Power=0.93), AVQI ($F[9]=3.12$, $p=0.002$, $\eta_p^2=0.14$, Observed Power=0.97), CAPE-V ($F[10]=2.58$, $p=0.006$, $\eta_p^2=0.12$, Observed Power=0.95), and VHI-30 scores ($F[10]=2.32$, $p=0.013$, $\eta_p^2=0.10$, Observed Power=0.93). For the VHI-10 subgroup, a significant effect was found for VHI-10 scores ($F[10]=2.26$, $p=0.014$, $\eta_p^2=0.05$, Observed Power=0.92) and a non-significant effect was found for MPT ($F[10]=0.72$, $p=0.701$, $\eta_p^2=0.04$, Observed Power=0.37), AVQI ($F[9]=0.90$, $p=0.527$, $\eta_p^2=0.21$, Observed Power=0.50), and CAPE-V ($F[10]=1.06$, $p=0.394$, $\eta_p^2=0.05$, Observed Power=0.55). The results of the post-hoc analyses were as follows:

VHI-30 Subgroup

The probability values for MPT scores were significant for the following disorder groups: VFI and transgender (p= <0.001), VFI and NSD (p= 0.001), VFI and MMLs (p= 0.009), atrophy/bowing and transgender (p= 0.009), VFI and neoplasm (p= 0.013), and MTD and transgender (p= 0.021).

For AVQI post-hoc analyses, the probability values for VFI and transgender, atrophy/bowing and transgender, and multifactorial and transgender were all significant at 0.002. Significant probability values at p=0.001 were also found for VFI and NSD, atrophy/bowing and NSD, as well as multifactorial and NSD for AVQI Post Hoc analyses.

For CAPE-V scores, probability values were significant for VFI and NSD (0.001), VFI and transgender (p= 0.002), and MTD and transgender (p= 0.003). Significant differences were also noted for NSD and MTD (p= 0.005), multifactorial and transgender (p= 0.009), atrophy/bowing and transgender (p=0.019), and neurological and transgender (p= 0.012).

For VHI-30 scores, post-hoc analysis indicated significant differences between the following disorder groups: VFI and “other” (0.012), VFI and multifactorial (p= 0.020), VFI and VCD (p= 0.021), VFI and NSD (p= 0.001), VFI and MMLs (p= 0.005), and VFI and atrophy/bowing (p= <0.001)

VHI-10 Subgroup

In this subgroup, it was found that VHI-10 scores were significantly different between VFI and NSD (p= 0.021), transgender and NSD (p= 0.024), and neurological and transgender (p= 0.012) disorder groups. The differences between VHI-10 scores for VFI and transgender (p= 0.001) and transgender and MMLs (p= 0.003) groups were also statistically significant.

4.3. Research Question 2: ***What is the strength of relationship between pre-treatment measures of the Acoustic Voice Quality Index (AVQI) and the Voice Handicap Index-30 (VHI-30)/Voice Handicap Index-10 (VHI-10)?***

Results indicated that the correlation between AVQI and VHI-30 scores was weak as defined by a negligible correlation[20]. The correlation between AVQI and VHI-10 scores was also considered to be weak. Both correlations were considered significant at the 0.01 level for a two-tailed test. Results are outlined in Table 6.

Table 6[21]

Pearson Correlation between VHI-30/VHI-10 and AVQI

	AVQI		
	n	p-value	Pearson's r
VHI-30	184	<0.001	0.272**
VHI-10	391	<0.001	0.287**

**. Correlation is significant at the 0.01 level (two-tailed).

[20] A negligible correlation is considered to range from 0.00 to 0.30 or 0.00 to -0.30, as defined by guidelines for interpreting Pearson's correlation coefficient results (Mukaka, 2012).
[21] Note: AVQI = Acoustic Voice Quality Index, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

4.4. Research Question 3: ***What is the strength of relationship between the pre-treatment measures of Maximum Phonation Time (MPT) and the VHI-30/VHI-10?***

Results indicated that there was a weak, negative correlation between MPT and VHI-30 scores, as well as MPT and VHI-10 scores[22]. The correlation between MPT and VHI-10 scores was considered significant at the 0.01 level for a two-tailed test. Results are outlined in Table 7.

Table 7[23]

Pearson Correlation between VHI-30/VHI-10 and MPT

		MPT	
	n	p-value	Pearson's r
VHI-30	211	0.100	-0.114
VHI-10	201	0.002	-0.212**

**. Correlation is significant at the 0.01 level (two-tailed).

[22] A weak correlation is defined as ranging from 0.00 to 0.30 or 0.00 to -0.30, indicating a negligible correlation per guidelines for interpreting Pearson's correlation coefficient results (Mukaka, 2012).
[23] Note: MPT = Maximum Phonation Time, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

4.5. Research Question 4: ***What is the strength of relationship between the pre-treatment perceptual measure of Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) and the VHI-30/VHI-10?***

Results indicated that the correlation between CAPE-V and VHI-30 scores was moderate, as defined by a low positive correlation[24]. There was also a moderate correlation between CAPE-V and VHI-10 scores. Both correlations were significant at the 0.01 level for a two-tailed test. Results are outlined in Table 8.

Table 8[25]

Pearson Correlation between VHI-30/VHI-10 and CAPE-V

	CAPE-V		
	n	p-value	Pearson's r
VHI	209	<0.001	0.457**
VHI-10	195	<0.001	0.401**

**. Correlation is significant at the 0.01 level (two-tailed).

[24] A low positive correlation is considered to range from 0.30 to 0.50, as defined by guidelines for interpreting Pearson's correlation coefficient results (Mukaka, 2012).
[25] Note: CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

4.6. Research Question 5: ***What is the relationship of measures AVQI, MPT, CAPE-V, and VHI-30/VHI-10 among different types of voice disorders?***

Bivariate correlation was performed to determine the correlation between measures of AVQI, MPT, CAPE-V, and VHI-30 scores based on the different disorder categories. As with previous analyses, data was split into VHI-30 and VHI-10 subgroups prior to analysis. Significance level was considered at alpha = 0.025 in order to address multiple testing and enhance the reliability and reproducibility of results. Results indicated that for the VHI-30 subgroup AVQI, MPT, CAPE-V, and VHI-30 scores were significantly related to each other in specific disorder categories: VFI, atrophy/bowing, multifactorial, NSD, MMLs, and MTD. For the VHI-10 subgroup, AVQI, MPT, CAPE-V, and VHI-10 scores were also significantly related to each other in specific disorder categories: VFI, atrophy/bowing, neurological, multifactorial, NSD, MMLs, and neoplasm. It should be noted that for the VHI-10 subgroup, AVQI, MPT, CAPE-V, and VHI-10 scores were all significantly related to one another within the MTD disorder type, likely due to small sample size.

VHI-30 Subgroup

A significant correlation was found between CAPE-V and AVQI scores for four disorder groups: VFI (r= 0.679), atrophy/bowing (r= -0.487), NSD (r= 0.518), and MMLs (r= 0.601). The multifactorial disorder group (r= 0.678) revealed a moderate correlation with CAPE-V and VHI-30 scores.

The NSD disorder group had a significant correlation between all measurements, ranging from weak (AVQI and VHI-30, r= 0.376) to moderate. Moderate correlations were demonstrated with MPT and AVQI (r= -0.438), MPT and CAPE-V (r= -0.467), AVQI and CAPE-V (r= 0.518), and CAPE-V and VHI-30 scores (r= 0.446).

MML disorder group demonstrated a significant moderate correlation between the AVQI and VHI-30 (r= 0.416), CAPE-V and VHI-30 (r= 0.560), as well as the CAPE-V and AVQI (r= 0.601). The MTD disorder group demonstrated a significant moderate correlation between AVQI and VHI-30 (r= 0.594), as well as CAPE-V and VHI-30 scores (r= -0.736). There were no significant correlations within the neurological disorder group, the "other" disorder group, the neoplasm disorder group, nor the transgender disorder group.

VHI-10 Subgroup

A significant correlation was found between MPT and CAPE-V (r= -0.659), MPT and AVQI (r= -0.603), and CAPE-V and AVQI (r= 0.633) for the VFI disorder group. All correlations within the VFI disorder group were considered significant and moderate. The neurological disorder group demonstrated a significant correlation for AVQI and VHI-10 (r= 0.837), CAPE-V and VHI-10 (r= 0.947), and CAPE-V and AVQI (r= 0.969). All correlations within the neurological disorder group are considered strong and significant.

The multifactorial disorder group also demonstrated a significant correlation for the aforementioned measures: AVQI and VHI-10 (r= 0.545), CAPE-V and VHI-10 (r= 0.698), and CAPE-V and AVQI (r= 0.857). The AVQI and VHI-10, as well as the CAPE-V and VHI-10 correlations, were considered significant and moderate whereas the CAPE-V and AVQI correlation was considered strong and significant.

Similar to the VHI-30 subgroup, the NSD disorder group had a significant correlation between all measurements except MPT & AVQI. The r-values for those measures are as follows: MPT and CAPE-V, r= -0.377; AVQI and CAPE-V, r= 0.470; AVQI and VHI-10, r= 0.214; CAPE-V and VHI-10, r= 0.449. For NSD, correlations for AVQI and CAPE-V as well as CAPE-V and VHI-10 were significant and moderate. Correlations were considered significant but weak for MPT and CAPE-V and AVQI and VHI-10 scores. MML disorder group demonstrated a significant and moderate correlation between CAPE-V and VHI-10 scores (r= 0.423). The neoplasm disorder group demonstrated a strong and significant correlation between MPT and CAPE-V scores (r= -0.996). There were no significant correlations within the "other" disorder group, nor the transgender disorder group.

Results are outlined in Table 9 (AVQI and CAPE-V), Table 10 (AVQI and MPT), Table 11 (AVQI and VHI), Table 12 (CAPE-V and MPT), Table 13 (CAPE-V and VHI), and Table 14 (MPT and VHI).

Table 9[26]

Pearson Correlation between AVQI and CAPE-V scores by disorder group

	AVQI and CAPE-V Scores					
	VHI-30			VHI-10		
	n	p-value	Pearson's r	n	p-value	Pearson's r
VFI	30	<0.001	0.679**	29	<0.001	0.633**
Atr/Bow	26	0.012	0.487*	4	0.322	0.678

[26] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, AVQI = Acoustic Voice Quality Index, CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

Neuro	3	0.663	0.505	6	0.001	0.969**
Other	6	0.129	0.690	13	0.090	0.489
Multi	18	0.214	0.308	14	<0.001	0.857**
Trans	6	0.498	-0.348	13	0.594	-0.163
NSD	39	0.001	0.518**	70	<0.001	0.470**
MMLs	36	<0.001	0.601**	28	0.072	0.345
MTD	12	0.064	0.551	2	-1.000**	-1.000**
Neo	5	0.295	0.590	4	0.091	0.909

**. Correlation is significant at the 0.01 level (two-tailed).

*. Correlation is significant at the 0.05 level (two-tailed)

Table 10[27]

Pearson Correlation between AVQI and MPT scores by disorder group

	AVQI and MPT Scores					
	VHI-30			VHI-10		
	n	p-value	Pearson's r	n	p-value	Pearson's r
VFI	31	0.083	-0.317	29	0.001	0.603**
Atr/Bow	26	0.636	0.097	4	0.793	-0.207
Neuro	3	0.411	-0.799	6	0.454	-0.383
Other	6	0.275	-0.534	13	0.508	-0.202

[27] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, AVQI = Acoustic Voice Quality Index, CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

Multi	18	0.931	-0.022	15	0.064	-0.490
Trans	7	0.580	0.256	13	0.432	-0.239
NSD	39	0.005	-0.438**	74	0.127	-0.179
MMLs	36	0.399	0.145	29	0.308	-0.196
MTD	12	0.215	0.386	2	-1.000**	-1.000**
Neo	5	0.982	-0.014	4	0.080	-0.920

**. Correlation is significant at the 0.01 level (two-tailed).

Table 11[28]

Pearson Correlation between AVQI and VHI scores by disorder group

	AVQI and VHI Scores					
	VHI-30			VHI-10		
	n	p-value	Pearson's r	n	p-value	Pearson's r
VFI	31	0.502	0.125	49	0.087	0.247
Atr/Bow	26	0.884	0.030	4	0.042	0.958*
Neuro	3	0.169	0.965	9	0.005	0.837**
Other	6	0.873	0.085	27	0.306	0.204
Multi	18	0.756	-0.079	32	0.001	0.545**
Trans	8	0.146	0.564	19	0.381	0.213
NSD	39	0.018	0.376*	186	0.003	0.214**

[28] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, AVQI = Acoustic Voice Quality Index, CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

MMLs	36	0.012	0.416*	55	0.056	0.259
MTD	12	0.042	0.594*	2	-1.000**	-1.000**
Neo	5	0.068	0.851	8	0.524	0.266

**. Correlation is significant at the 0.01 level (two-tailed).

*. Correlation is significant at the 0.05 level (two-tailed)

Table 12[29]

Pearson Correlation between CAPE-V and MPT scores by disorder group

	CAPE-V and MPT Scores					
	VHI-30			VHI-10		
	n	p-value	Pearson's r	n	p-value	Pearson's r
VFI	32	0.366	-0.165	31	<0.001	-0.659**
Atr/Bow	27	0.189	-0.261	4	0.241	0.759
Neuro	3	0.252	-0.923	6	0.491	-0.354
Other	11	0.275	-0.361	13	0.427	-0.241
Multi	21	0.475	-0.165	15	0.125	-0.414
Trans	6	0.212	0.596	15	0.076	0.472
NSD	47	0.001	0.467**	75	0.001	-0.377**
MMLs	42	0.824	-0.035	31	0.526	-0.118
MTD	15	0.706	0.106	2	-1.000**	-1.000**

[29] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, AVQI = Acoustic Voice Quality Index, CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

Neo	5	0.915	0.066	4	0.004	-0.996**

**. Correlation is significant at the 0.01 level (two-tailed).

Table 13[30]

Pearson Correlation between CAPE-V and VHI scores by disorder group

	CAPE-V and VHI Scores					
	VHI-30			VHI-10		
	n	p-value	Pearson's r	n	p-value	Pearson's r
VFI	32	0.192	0.237	31	0.066	0.334
Atr/Bow	27	0.386	0.174	4	0.518	0.482
Neuro	3	0.494	0.714	6	0.004	0.947**
Other	11	0.909	0.039	13	0.560	0.179
Multi	21	0.001	0.678**	15	0.004	0.698**
Trans	6	0.628	0.254	15	0.343	0.263
NSD	47	0.002	0.446**	74	<0.001	0.449**
MMLs	42	<0.001	0.560**	31	0.018	0.423*
MTD	15	0.002	0.736**	2	-1.000**	-1.000**
Neo	5	0.114	0.787	4	0.693	0.307

**. Correlation is significant at the 0.01 level (two-tailed).

*. Correlation is significant at the 0.05 level (two-tailed)

[30] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, AVQI = Acoustic Voice Quality Index, CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

Table 14[31]

Pearson Correlation between MPT and VHI scores by disorder group

	MPT and VHI Scores					
	VHI-30			VHI-10		
	n	p-value	Pearson's r	n	p-value	Pearson's r
VFI	33	0.092	-0.299	31	0.726	-0.066
Atr/Bow	27	0.807	0.049	4	0.910	-0.090
Neuro	3	0.242	-0.929	6	0.507	-0.342
Other	11	0.402	-0.281	13	0.348	-0.284
Multi	21	0.196	-.294	16	0.051	-0.496
Trans	7	0.643	0.215	15	0.513	0.183
NSD	47	0.335	-0.144	78	0.153	-0.163
MMLs	42	0.359	0.145	32	0.226	-0.220
MTD	15	0.237	0.325	2	-1.000**	-1.000**
Neo	5	0.568	-0.347	4	0.690	-0.310

**. Correlation is significant at the 0.01 level (two-tailed).

*. Correlation is significant at the 0.05 level (two-tailed)

[31] Note: VFI = vocal fold insufficiency, Atr/Bow = atrophy or bowing, Neuro = neurological, Multi = multifactorial, Trans = transgender, NSD = non-specific dysphonia, VCD = vocal cord dysfunction, MMLs = mid-membranous lesions, MTD = muscle tension dysphonia, Neo = neoplasm, AVQI = Acoustic Voice Quality Index, CAPE-V = Consensus Auditory-Perceptual Evaluation of Voice, VHI-30 = Voice Handicap Index-30, VHI-10 = Voice Handicap Index-10

CHAPTER V: DISCUSSION

The goal of this study was to further understand the relationship between the VHI-30/VHI-10, and acoustic, aerodynamic, and auditory-perceptual measures of voice. Further knowledge of these relationships should expand clinical understanding which can aid in prioritizing intervention and increase the likelihood of improving current parameters of voice with the greatest influence on the patient's quality of life. In addition, further understanding could contribute to current clinical knowledge to facilitate the selection of treatment methods for specific patients.

Research Question 1: What is the effect of disorder type on the measurement of Maximum Phonation Time (MPT), Acoustic Voice Quality Index (AVQI), Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), and Voice Handicap Index scores (VHI – VHI-30 version and VHI-10 version)?

Results indicated that for the subgroup administered the VHI-30, there was a significant effect of disorder on MPT, CAPE-V, AVQI, and VHI-30 scores. For the VHI-10 subgroup, there was a significant effect of disorder only on VHI-10 scores. Previous findings by Rosen et al. (2004) argued that the VHI-10 may be more sensitive to detecting differences in diagnoses (Rosen, Lee, Osborne, Zullo, & Murry, 2004), however, the results of this study present potential rebuttals to this argument. It should also be noted that the VHI-30 was the only instrument on voice-related measures to meet the criteria for reliability, validity, and availability of normative data for the Agency for Health Care Research and Quality in 2002. While this was determined prior to the development of the VHI-10, it is important to note that the VHI-10 has yet to be as rigorously tested in the same regard as the VHI-30 (Rosen, Lee, Osborne, Zullo, & Murry, 2004). Results of this study indicate that the VHI-30 may be more sensitive to the overall effect that disorder has on the patient and their individual voice measures obtained pre-treatment. Future studies may wish to examine the effect of disorder on the VHI-30 and VHI-10 comparatively among larger sample sizes, potentially with consideration of clinical voice measurements similar to those within this study.

Research Questions 2-3: What is the strength of relationship between pre-treatment measures of the Acoustic Voice Quality Index (AVQI) and Maximum Phonation Time (MPT) with the Voice Handicap Index (VHI-30 and VHI-10)?

There was a statistically significant but weak relationship between pre-treatment measures of the AVQI for both VHI-30 and VHI-10, as well as MPT for both VHI-30 and VHI-10. Results are consistent with previous literature that found weak correlations between patient-reported measures and clinical measures of voice (Gillespie, Gooding, Rosen, & Gartner-Schmidt, 2014; Karlsen, Sandvik, Heimdal, & Aartad, 2020; Lopes, da Silva, Simoes, Evangelista, Silva, Almeida, & de Lima-Silva, 2017; Lopes, da Silva, Evangelista, da Silva, Simoes, Silva, de Lima-Silva, & de Almeida, 2016). Results indicate that while there may be some correlation between AVQI and MPT with VHI-30/VHI-10 scores, overall, the strength of this correlation should be considered when examining pre-treatment measures. The correlation between these two measures may be weak due to the difference between the patient's experience and interpretation of their voice as compared to a clinician's training to quantify voice. Subjectively, patients are able to report and rate the severity of their voice disorder; however, this may not always align with what is determined objectively when acoustic and aerodynamic measures are performed to quantify that severity, ultimately resulting in a weak correlation between the two. Overall, the results of this study are consistent with what has previously been reported in regard to the relationship between acoustic and aerodynamic measures with patient-reported outcome measures. Future studies may wish to examine how demographics such as age, gender, and occupation affect this correlation.

Research Question 4: What is the strength of relationship between the pre-treatment perceptual measure of Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) and the VHI-30/VHI-10?

There was a significant and moderate relationship between CAPE-V and VHI-30 scores, as well as CAPE-V and VHI-10 scores. Findings are consistent with previous studies that have found relationships among auditory-perceptual ratings, particularly for the perceptual correlates of vocal functioning such as pitch, roughness, and strain (Houle & Johnson, 2021; Karlsen, Sandvik, Heimdal, & Aartad, 2020; Lopes, da Silva, Simoes, Evangelista, Silva, Almeida, & de Lima-Silva, 2017). Results may be indicative that subjective measures of voice are better related to one another as compared to objective measures of voice. This is because the current study found a significant, but weak, relationship between VHI-30/VHI-10 and AVQI as well as VHI-30/VHI-10 and MPT whereas a significant and moderate relationship was found between subjective measures (CAPE-V and VHI-30/VHI-10). While it was hypothesized that subjective measurements would not correlate with one another, overall results indicate that subjective measurements of voice are related and may give a better picture of understanding the severity and impact of the voice disorder on a patient as compared to that of objective measurements (AVQI and MPT).

Research Question 5: What is relationship of measures AVQI, MPT, CAPE-V, and VHI-30/VHI-10 among different types of voice disorders?

There was a significant and moderate relationship between disorder and CAPE-V, disorder and AVQI scores, as well as disorder and VHI-30 scores for the VHI-30 subgroup.

There was a weak relationship between disorder and MPT within this subgroup. For the VHI-10 subgroup, there was a moderate relationship found between disorder and measures of MPT, AVQI, CAPE-V, and VHI-10 scores. Results indicate that while disorder had an effect on all measures within both the VHI-30 and VHI-10 subgroups, there appears to be a greater relationship between disorder and measures within the VHI-10 subgroup as compared to disorder and measures within the VHI-30 subgroup; the primary difference being a weak correlation of disorder with MPT (VHI-30 subgroup) as compared to a moderate correlation of disorder and MPT (VHI-10 subgroup). Results indicate that future clinical applications should consider MPT scores in particular when comparing other measures of voice to patients who have completed the VHI-30 or the VHI-10. Disorder groups with the greatest relationship with measurements of voice and possible explanations are outlined as follows.

VHI-30 Subgroup

For the VHI-30, the disorder groups that had the greatest relationship with measurements of voice were non-specific dysphonia (NSD), mid-membranous lesions (MMLs), and muscle tension dysphonia (MTD). Interestingly, NSD correlated with all measures. This may be due to the fact that NSD is dysphonia unspecified, where there is no underlying cause that can be pinpointed, and a lack of reasoning for dysphonia may cause an overall decrease in quality of life and increase in severity. It should also be noted, however, that NSD was the largest disorder group within the study.

MMLs may have a greater relationship with measurements of voice due to the extent to which this disorder group encompasses; for example, individuals categorized into the MMLs disorder group were diagnosed with a variety of disorders such as granulomas, vocal fold nodules, polyps, scars, cysts, sulci, and Reinke's edema. The

MTD disorder group had a correlation within subjective measurements (VHI-30 and CAPE-V), which aligns with what is known about MTD and its impact on functionality (Behlau, Madazio, & Oliveria, 2015). This correlation is likely due to the fact that subjective measurements of voice were designed to determine the impact of a voice disorder on daily functioning. The relationship of MTD with VHI-30 and AVQI may be due to an individual's awareness of how their voice sounds to listeners and its severity, likely reflected in their VHI-30 score.

VHI-10 Subgroup

Within the VHI-10 subgroup, the disorder groups that had the greatest relationship with measurements of voice were NSD, multifactorial, neurological, and vocal fold insufficiency (VFI). Similar to the VHI-30 subgroup, NSD had correlations with many of the measurements; the only difference is that NSD did not correlate with objective measurements (MPT and AVQI). This could be due to considerations of psychogenic dysphonia, the loss of voice in the absence of apparent structural or neurological pathology. If there are no anatomical or physiological impairments, then it is possible that acoustic and aerodynamic, objective, measurements of voice would not be indicative of a disorder (Sudhir, Chandra, Shivashankar, & Yamini, 2009).

Multifactorial and neurological disorder groups both had a relationship with the same measurements: CAPE-V and AVQI, VHI-10 and AVQI, and CAPE-V and VHI-10. Participants were assigned to the multifactorial disorder group if they had concomitant diagnoses, such as dysphagia, dysarthria, reflux, and multiple diagnoses of voice disorders. The neurological group consisted of participants' diagnoses with laryngeal spasm, dyskinesia, or tremor either in isolation or within the presence of Parkinsonism.

Both of these disorder groups may have had a relationship with the aforementioned measurements due to their overall impact on quality of life, affecting both CAPE-V and VHI-10 outcomes. It should be noted that these disorder groups did not have a relationship with any correlations containing aerodynamic measurements; this could be due to the sensitivity of the AVQI to analyze acoustic parameters of the voice that may affect the physiological parameters of voice production rather than the anatomical components (Maryn, De Bodt, & Roy, 2009).

VFI, consisting of vocal fold paresis (weakness) and paralysis (immobility), also had an interesting relationship with measurements of voice. For example, this disorder group had an effect on both CAPE-V and AVQI, and CAPE-V and MPT, likely due to the significant impact VFI can have on auditory-perceptual measures of voice and the clinician's interpretation of severity. Interestingly, this disorder group was also the only group to have a significant correlation with MPT and AVQI. This is potentially due to the impact this disorder group has on the anatomy and physiology of the vocal folds, such as decreased breath support likely resulting in a decrease in MPT, and the potential for diplophonia, the production of two separate tones, which would significantly impact AVQI scores (Hong & Kim, 1999).

In regard to the implications of study results on current understanding and clinical application, it should be noted that disorder appears to have a significant correlation with VHI-30 scores and measures of voice (aerodynamic, acoustic, auditory-perceptual) as compared to the VHI-10. The reasoning for this could be due to a larger quantity of questions within the VHI-30, the population in which the VHI-10 was normed on, and the overall development of the VHI-10 as compared to the VHI-30 (Rosen, Lee, Osborne, Zullo, & Murry, 2004). Results also indicate that the objective measures of voice utilized in this study may assess different constructs of vocal function, as they were not strongly correlated with each other. However, subjective measures such as the VHI-30 and the CAPE-V were more strongly associated.

CONCLUSION

The results of this study found that objective measures of voice did not correlate strongly with patient perceptions of voice handicap, while subjective measures of voice quality were more strongly associated with the patient's perception of voice handicap. It is also important to note that disorder type affected clinical measures for the VHI-30 subgroup to a greater extent than the subgroup who completed the VHI-10. This may support the notion that the VHI-30 is more sensitive to different voice disorder diagnoses than the VHI-10 version.

Future studies may wish to further examine the strength of the relationship between MPT, AVQI, and CAPE-V for both the VHI-30 and the VHI-10, particularly in other languages. Another consideration may be to examine the effect of demographics such as age, gender, and occupation on the correlation between these measures. This consideration combined with variable results of prior studies (Van Houtte, Van Lirder, D'Haeseleer, & Claeys, 2010; Coyle, Weinrich, & Stemple, 2001) supports a hypothesis suggested by Watanabe, Sato, Honkura, Kawamoto-Hirano, Kashima, and Katori (2019) that objective parameters of voice may correlate with the VHI-30 and VHI-10 differently for a homogenous patient population.

REFERENCES

Behlau, M., Madazio, G., & Oliveira, G. (2015). Functional dysphonia: Strategies to improve patient outcomes. Patient Related Outcome Measures, 6, 243-253. https://doi.org/10.2147/PROM.S68631

Cheng, J., & Woo, P. (2010). Correlation between the voice handicap index and voice laboratory measurements after phonosurgery. *Ear, Nose, & Throat Journal*, *89*(4), 183–188.

Coyle, S. M., Weinrich, B. D., & Stemple, J. C. (2001). Shifts in relative prevalence of laryngeal pathology in a treatment-seeking population. Journal of Voice, 15(3), 424-440. https://doi.org/10.1016/S0892-1997(01)00043-1

Gillespie, A. I., Gooding, W., Rosen, C., & Gartner-Schmidt, J. (2014). Correlation of VHI-10 to voice laboratory measurements across five common voice disorders. *Journal of Voice*, 28(4), 440-448. https://doi.org/10.1016/j.jvoice.2013.10.023

Hapner, E. R. & Johns, M. (2021). *Laryngeal Videostroboscopy: Part One & Part Two* [MedBridge presentation]. https://www.medbridgeeducation.com/courses/details/laryngeal-videostroboscopy-part-one-edie-hapner-michael-johns-speech-language-pathology

Hong, K. H. & Kim, H. K. (1999). Diplophonia in unilateral vocal fold paralysis and intracordal cyst. *Official Journal of American Academy of Otolaryngology – Head and Neck Surgery*, 121(6), 815-819. https://doi.org/10.1053/hn.1999.v121.a94214

Houle, N. & Johnson, A. M. (2021). Relationships across clinical measures of vocal quality and functioning and their relationship with patient perception. *American Journal of Speech-Language Pathology*, 30(6), 2668-2680. https://doi.org/10.1044/2021_AJSLP-21-00128

Hsiung, M. W., Pai, L., & Wang, H. W. (2002). Correlation between voice handicap index and voice laboratory measurements in dysphonic patients. *European Archives of Oto-Rhino-Laryngology: Official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS)*, 259(2), 97–99. https://doi.org/10.1007/s004050100405

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The voice handicap index (VHI): Development and validation. *American Journal of Speech-Language Pathology*, 6(3), 66-70. https://doi.org/10.1044/1058-0360.0603.66

Karlsen, T., Sandvik, l., Heimdal, J. H., & Aartad, H. J. (2020). Acoustic voice analysis and maximum phonation time in relation to voice handicap index score and larynx disease. *Journal of Voice*, 34(1), 161.e27-161.e35. https://doi.org/10.1016/j.jvoice.2018.07.002

Lopes, L. W., da Silva, H. F., Evangelista, D. D. S., da Silva, J. D., Simoes, L. B., Silva, P. O. C. E., de Lima-Silva, M. F. B., & de Almeida, A. A. F. (2016). Relationship between vocal symptoms, severity of voice disorders, and laryngeal diagnosis in patients with voice disorders. *CODAS*, 28(4), 439-445. https://doi.org/10.1590/2317-1782/20162015062

Lopes, L. W., da Silva, J. D., Simoes, L. B., Evangelista, D. D. S., Silva, P. O. C., Almeida, A. A., & de Lima-Silva, M. F. B. (2017). Relationship between acoustic measurements and self-evaluation in patients with voice disorders. Journal of Voice, 31(1), 119.e1-119.e10. https://doi.org/10.1016/j.voice.2016.02.021

Maryn, Y., De Bodt, M., & Roy, N. (2009). The acoustic voice quality index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*, 43(3), 161-174. https://doi.org/10.1016/j.jcomdis.2009.12.004

Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal: The Journal of Medical Association of Malawi*, 24(3), 69-71.

Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the voice handicap index-10. *The Laryngoscope*, 114(9), 1549-1556. https://doi.org/10.1097/00005537-200409000-00009

Rosen, C. A., Murry, T., Zinn, A., Zullo, T., & Sonbolian, M. (2000). Voice Handicap Index change following treatment of voice disorders. *Journal of Voice*, 14(4), 619-623. https://doi.org/10.1016/s0892-1997(00)80017-x

Sudhir, P. M., Chandra, P. S., Shivashankar, N., & Yamini, B. K. (2009). Comprehensive management of psychogenic dysphonia: A case illustration. *Journal of Communication Disorders*, 42(5), 305-312. https://doi.org/10.1016/j.jcomdis.2009.04.003

Thijs, Z., Knickerbocker, K., & Watts, C. R. (2020). Epidemiological patterns and treatment outcomes in a private practice community voice clinic. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2020.06.025

Van Houtte, E., Van Lirder, K., D'Haeseleer, E., & Claeys, S. (2010). The prevalence of laryngeal pathology in a treatment-seeking population with dysphonia. *The Laryngoscope,* 120(2), 306-312. https://doi.org/10.1002/lary.20696

Watanabe, K., Sato, T., Honkura, Y., Kawamoto-Hirano, A., Kashima, K., & Katori, Y. (2019). Characteristics of the voice handicap index for patients with unilateral vocal fold paralysis who underwent arytenoid adduction. *Journal of Voice*, 34(4), 649.e1-649.36. https://doi.org/10.1016/j.voice.2005.03.006

Watts, C. R. & Awan, S. N. (2019). *Laryngeal function and voice disorders: Basic science to clinical practice*. Thieme.

Watts, C. R. & Knickerbocker, K. (2019). Characteristics of a treatment-seeking population in a private practice community voice clinic: An epidemiological study. *Journal of Voice*, 33(4), 429-434. https://doi.org/10.1016/j.jvoice. 2017.11.019

Wheeler, K. M., Collins, S. P., & Sapienza, C. M. (2006). The relationship between VHI scores and specific acoustic measures of mildly disordered voice production. *Journal of Voice*, 20(2), 308-317. https://doi.org/10.1016/j.jvoice.2005.03.006

Zhao, E. E., Nguyen, S. A., Salvador, C. D., & O'Rourke, A. K. (2020). A meta-analysis of the association between the voice handicap index and objective voice analysis. *Journal of Speech, Language, and Hearing Research*, 63(10), 3461–3471. https://doi.org/10.1044/2020_JSLHR-20-00209

ABSTRACT

Purpose: Historically research has found no significant correlation between clinical measures of voice and patient-perceived voice handicap, or significant but weak correlations between the two. This study aimed to further expand the clinical understanding of the relationship between acoustic, aerodynamic, and auditory-perceptual measures of voice to patient-perceived voice

Methods: The following data were extracted from an existing database from a private practice community voice clinic: (1) auditory perceptual ratings of voice quality (Consensus Auditory-Perceptual Evaluation of Voice, CAPE-V), (2) maximum phonation time (MPT), (3) Acoustic Voice Quality Index (AVQI) scores, and (4) Voice Handicap Index scores (VHI-30 or VHI-10). All data represent measurements at pre-treatment, prior to the application of any voice therapy. Because part of the sample was administered the VHI-30 before the practice switched to administering the VHI-10, data was split into VHI-30 and VHI-10 subgroups for all data analyses. ANOVAs were applied to the data set to determine the effect of disorder on each measurement, and bivariate correlation was also performed in order to determine the correlation between pre-treatment measures of voice (AVQI, MPT, CAPE-V) and patient-perceived voice handicap (VHI-30, VHI-10).

Results: The results of this study found that objective measures of voice did not correlate strongly with patient perceptions of voice handicap, while subjective measures of voice quality were more strongly associated with the patient's perception of voice handicap. Results also indicated that disorder type affected clinical measures for the VHI-30 subgroup to a greater extent than the subgroup who completed the VHI-10.

Conclusion: Future studies may wish to further examine the strength of the relationship between MPT, AVQI, and CAPE-V for both the VHI-30 and the VHI-10, particularly in other languages. Future studies may also examine the effect of demographics such as age, gender, and occupation on the correlation between these measures.