This study examined the impact of using video modeling techniques on the efficiency and stability of measurements during acoustic voice assessment. Thirty-four participants between the ages of 18 and 30 years were recruited for participation in the study and randomly assigned to either the treatment or control group for comparative analysis. Experimental group participants watched a training video prior to receiving verbal instruction and completing acoustic voice assessment tasks. Control group participants received the same conditions with the exception of watching a training video. Findings from independent samples t-tests suggest that there is not a significant difference between the amount of time, number of cues, mean values, and standard deviations across three trials when comparing those who have watched a training video and those who have not. However, a trend towards more stable results was seen on maximum effort tasks.
IMPACT OF VIDEO MODELING TECHNIQUES ON EFFICIENCY AND EFFECTIVENESS OF CLINICAL VOICE ASSESSMENT

A Thesis

Submitted to the

Faculty of Miami University

in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Speech Pathology and Audiology

by

Samantha Lauren Bowyer

Miami University

Oxford, Ohio

2014

Advisor_________________________

Susan Baker Brehm, Ph.D.

Reader_________________________

Barbara Weinrich, Ph.D.

Reader_________________________

Wendy LeBorgne, Ph.D.
# TABLE OF CONTENTS

CHAPTER I (INTRODUCTION AND REVIEW OF LITERATURE) ........................................1
   Productivity and Effectiveness in Speech-Language Pathology ............................1
      Productivity  ....................................................................................................1
      Effectiveness  ..................................................................................................1
   Technology in Speech-Language Pathology .........................................................2
   Voice Disorders  .................................................................................................3
   Acoustic Voice Assessment  .................................................................................4
   Statement of the Problem/Purpose of the Study ................................................5

CHAPTER II (METHODS) ..........................................................................................7
   Participants  .........................................................................................................7
      Criteria for Participation  ..................................................................................7
   Procedure  ..........................................................................................................7
      Informed Consent  .............................................................................................8
      Questionnaire  ..................................................................................................8
      Hearing Screening  ............................................................................................8
      Voice Screening  ..............................................................................................8
      Testing Protocol: Control Group .................................................................9
      Testing Protocol: Experimental Group .......................................................9
      Acoustic Measurements  ..................................................................................9
   Analysis of Data  ................................................................................................10

CHAPTER III (RESULTS) ........................................................................................11
   Acoustic Assessment Completion Time ..........................................................11
   Number of Cues  .................................................................................................11
   Acoustic Measurement Means ........................................................................12
      Fundamental Frequency  ...............................................................................12
      Intensity  ..........................................................................................................13
      Speaking Fundamental Frequency ...............................................................13
      Speaking Intensity  ..........................................................................................14
      Maximum Phonation Time .............................................................................15
LIST OF FIGURES

Figure 1 – Mean Instructional Time ................................................................. 11
Figure 2 – Mean Number of Cues ................................................................. 12
Figure 3 – Mean Fundamental Frequency During Sustained /a/ Task ........ 12
Figure 4 – Mean Intensity During Sustained /a/ Task ................................. 13
Figure 5 – Mean Speaking Fundamental Frequency .................................. 14
Figure 6 – Mean Speaking Intensity ............................................................. 14
Figure 7 – Mean Maximum Phonation Time .............................................. 15
Figure 8 – Mean Maximum Frequency ....................................................... 16
Figure 9 – Mean Minimum Frequency ......................................................... 16
Figure 10 – Standard Deviation of Fundamental Frequency Measurements During Sustained /a/ Task ................................................................. 17
Figure 11 – Standard Deviation of Intensity Measurements During Sustained /a/ Task .............. 18
Figure 12 – Standard Deviation of Speaking Fundamental Frequency Measurements .......................... 18
Figure 13 – Standard Deviation of Speaking Intensity ................................. 19
Figure 14 – Standard Deviation of Maximum Phonation Time Measurements ........................................ 20
Figure 15 – Standard Deviation of Maximum Frequency Measurements .................. 20
Figure 16 – Standard Deviation of Minimum Frequency Measurements .............. 21
ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Brehm, for her research mentorship throughout my graduate career. The knowledge and guidance that she has provided over the past two years has been exceptional and greatly appreciated. I would also like to thank the members of my graduate committee, Dr. Weinrich and Dr. LeBorgne. Finally, I would like to thank my family and friends, Mom, Dad, Andy, Mike, Lauren, Jenny, and Amelia, for their encouragement and love during this project and throughout my graduate studies.
CHAPTER I
Introduction and Review of Literature

With climbing health care costs in recent years, efficiency has become increasingly important in all aspects of medical practice. Between 1990 and 2010, the national health expenditure grew from 724.3 billion dollars to over 2.59 trillion dollars (Centers for Disease Control [CDC], 2012). At a time when emphasis is placed on cutting governmental spending, the importance of finding ways to intersect effective healthcare and efficient practice is brought to the forefront of discussions throughout the medical community. The passing of the Affordable Care Act, which has contributed to substantial evolution of the United States healthcare system, further emphasizes these ideals. In fact, Title III of the Affordable Care Act is dedicated to making improvements in the quality and efficiency of healthcare in the United States (PL 111-148, 2010). In many subfields of medicine, including speech-language pathology, growth of technology has offered unique opportunities to bring effective and efficient clinical practice to fruition.

Productivity and Effectiveness in Speech-Language Pathology

Productivity. Productivity is a constant concern for practicing speech-language pathologists. Productivity may be defined as the number of hours in direct patient care divided by the total number of hours worked (American Speech-Language-Hearing Association [ASHA], 2013). Thus, productivity is a measurement of efficiency in the workplace (Dennis & Gozenbach, 2011). In the 2009 Speech-Language Pathology (SLP) Health Care Survey, 59 percent of clinicians surveyed indicated that their place of employment had requirements for productivity (ASHA, 2009). Requirements regarding productivity in the workplace necessitate providing effective care in the most efficient manner possible. A variety of factors may affect reported productivity in a given healthcare setting. Such factors include the complexity of services provided, staffing, availability of space and equipment, work duties of practitioners not related to direct patient care, coding systems, and labor mapping (Dennis & Gozenbach, 2011).

Effectiveness. While a focus on efficient, productive practice is important, focus must also be placed on effective practice. The Principle of Ethics I in the ASHA Code of Ethics states, “Individuals shall honor their responsibility to hold paramount the welfare of persons they serve professionally” (ASHA, 2010). Rule of Ethics I under this principle states, “Individuals shall evaluate the effectiveness of services rendered and of products dispensed” (ASHA, 2010).
Taking both of these statements into consideration, clinicians must not forgo effective practice for the sake of efficiency. The challenge facing speech-language pathologists lies in balancing efficiency and effectiveness to achieve clinical practice that is both productive and beneficial to patients being served.

**Technology in Speech-Language Pathology**

The field of speech-language pathology has evolved extensively with the onslaught of digital technologies. Using telecommunications technology, speech-language pathologists can effectively treat patients with limited accessibility to services from a remote location (Mashima et al., 2003). Software that can be used by practitioners to monitor and modify computer-based home therapy programs from remote locations has been found to increase picture naming abilities, confidence, and communication participation of patients with aphasia (Mortley, Wade, & Enderby, 2004). The intervention was deemed acceptable by patients when paired with home visits for assessments or technology trouble-shooting and phone calls (Mortley et al., 2004).

Methods for augmentative and alternative communication (AAC) now include the use of “high tech” devices such as iPads, iPods, and dedicated speech devices. Cognitive prostheses, such as personal digital assistants, tablets, and smartphones, may be provided to patients during cognitive-linguistic therapy to compensate for memory, attention, and organizational deficits (de Domingo, 2013). Digital data storage and review is feasible through equipment such as the KayPentax Video Viewer software, which allows videos created using Kay’s Digital Strobe (KDS) and Digital Swallowing Workstation (DSW) to be stored and reviewed on a remote computer (KayPentax, n.d.). The use of electronic health records in all aspects of the medical profession has been found to be more efficient and cost-effective than paper medical charts (HealthIT.gov, n.d.). These examples reflect some of the many ways in which digital technologies are used in speech-language pathology.

Video technology for modeling purposes has been used extensively in speech-language pathology as an intervention for children and adults with Autism Spectrum Disorders (ASD) and other developmental disorders. Literature surrounding video modeling with individuals with ASD examines video modeling as a method of teaching conversational skills, play skills, pro-social behavior, and workplace tasks. In a comparative study examining video modeling and live modeling for teaching children with ASD a variety of social and cognitive-linguistic tasks, findings suggested that video modeling resulted in more efficient skill acquisition than live
modeling for four of the five children in the study (Charlop-Christy, Le, & Freeman, 2000). A similar study yielded conflicting evidence regarding the superiority of video modeling, suggesting that some children with ASD learn better from a live model or a combination of video modeling and live modeling; however, findings suggested that participants demonstrated greater visual attention towards video models compared to live models (Wilson, 2013).

The efficiency of video modeling is discussed extensively in the literature regarding treatment interventions for ASD. For example, video models can be viewed multiple times until the target individual understands the task being presented. Live modeling, on the other hand, is less accurate in additional presentations and requires time from the individual presenting the model (Ayres & Lagone, 2005). Similarly, a single video can be used with multiple individuals, further reducing the amount of time a clinician spends providing a model (McCoy & Hermansen, 2007). In addition, some have suggested that video modeling is more motivating than live modeling, in that video modeling is considered a novel addition to a learning environment (Dowrick, 1986).

Video modeling techniques have also been used to enhance voice therapy. One such way is by providing adults enrolled in a voice therapy program with models of home practice exercises on a portable digital media player (van Leer & Connor, 2012). Patients were found to be more compliant in completing voice exercise regimens when provided with a video model in addition to written instructions than when only given written instructions (van Leer & Connor, 2012). Statistically significant differences in practice frequency and motivation to complete home voice exercises were found in favor of the video model paired with written instructions (van Leer & Connor, 2012).

**Voice Disorders**

Voice therapy programs, as described above, are designed for individuals with voice disorders. Voice disorders occur when the perceptual aspects of voice, such as quality, pitch, and loudness, differ from the normal limits for such properties. Voice disorders may also occur when the vocal mechanism no longer meets the needs of the speaker in some way (Stemple, Glaze, & Klaben, 2010). Voice pathologies are numerous and vary greatly depending on etiologic correlates. For example, voice disorders may be organic, caused by structural problems of the vocal tract, or they may be neurologic, meaning innervations and muscle control involved in one or more components of voice production is impaired (Boone, McFarlane, & Von Berg,
2005). Additionally, a large group of voice disorders exist that lack structural or neurologic correlates, but rather are caused by vocal misuse (Boone et al., 2005). For example, such disorders, known as functional voice disorders, may be caused by vocally traumatic behaviors, such as shouting, coughing, throat clearing, speaking at an inappropriate pitch, and loud talking (Stemple et al., 2010). Vocally traumatic behaviors cause laryngeal hyperfunction, forcing the vocal folds together unnecessarily. The most common laryngeal pathologies in adults include nodules, polyps, edema, cancer, and vocal fold paralysis (Coyle, Weinrich, & Stemple, 2001).

Although data regarding the frequency of voice disorders occurrence varies considerably, one study suggests that lifetime prevalence is upwards of 30 percent (Roy, Merrill, Gray, & Smith, 2005). Individuals working in occupations that require consistent vocal use, such as teachers, vocal performers, telemarketing, etc. are more at risk for developing voice disorders. Furthermore, individuals who consistently talk over noise are also at risk (Stemple et al., 2010). As can be seen by examining the etiologic correlates, while vocal performers comprise a defined population of individuals who develop voice disorders, a large group of patients who are treated for voice disorders are not experienced with performing voice tasks, such as those asked of them during clinical voice assessment, which are described below.

**Acoustic Voice Assessment**

Acoustic voice assessment is a key component of voice evaluation, in that it provides supplemental information that cannot be obtained from the case history, laryngoscopic examination of the vocal folds, aerodynamic assessment, and clinical observations (Stemple et al., 2010). Acoustic measurements provide clinicians with objective data to compare with normative values, aiding in the diagnosis of voice disorders. Such measurements also provide a baseline for therapy. Progress can be monitored by comparing measurements obtained during assessment with measurements taken periodically during the course of therapy or at the conclusion of a course of therapy. Acoustic measurements include fundamental frequency, intensity, phonation range, and maximum phonation time. Fundamental frequency is the rate at which the vocal folds most naturally vibrate in a given individual. Fundamental frequency is perceived as an individual’s pitch (Baken, & Orlikoff, 2000). Intensity relates to the sound pressure level of an individual’s voice, which is perceived as an individual’s loudness (Baken & Orlikoff, 2000). Phonation range is the measurement of the highest and lowest frequencies that
an individual is capable of producing. Maximum phonation time is the amount of time that an individual can sustain a vowel after taking a maximum inhalation (Stemple et al., 2010).

Tasks performed to achieve measurements of the aforementioned components of acoustic voice assessment are typically novel to patients. Fundamental frequency and intensity measurements are generally taken from samples of running speech and sustained vowels. Sustained vowels are also used to measure maximum phonation time. During the maximum phonation time task, a patient would be asked to sustain an /a/ for as long as possible on a single breath. To measure phonation range, patients would be instructed to glide to their highest note on “whoop,” starting at a comfortable pitch, followed by a glide to their lowest note on “boom.” Most individuals do not have experience with these types of tasks; thus, explicit instruction of how each task is performed is necessary for patient success in task performance.

Acoustic measurements are typically taken over multiple trials. Classic literature in the field of speech pathology suggests a practice effect on maximum performance tests (for example, maximum phonation time) leads to improvement across trials (Kent, Kent, & Rosenbek, 1987). Improvements may be seen in as many as 10 to 15 trials, but researchers suggest that with careful instruction and practice trials, adequately stable measurements may be obtained in three trials (Kent et al., 1987).

Statement of the Problem/Purpose of the Study

The melding of efficient and effective clinical practice is of utmost importance as the United States faces an evolution of the healthcare system under the Affordable Care Act. As evidenced through growth of technology in a variety of aspects of speech-language pathology, a mode of accomplishing more efficient and effective clinical voice assessment may be to examine how innovative technology can enhance clinical practice. Principles of video modeling have been used successfully in speech and language settings with a variety of populations, including individuals with ASD and adults with voice disorders completing home exercise programs (Charlop-Christy et al., 2000; Wilson, 2013; van Leer & Connor, 2012). Video modeling for voice assessment tasks, however, has not been explored. The possibility of increasing the efficiency and effectiveness of completing voice assessment through the use of video modeling on a portable tablet device was investigated in the present study. Healthy adults without any known pathology of the larynx were provided with either a video model on a digital media player prior to voice assessment paired with verbal instructions or only verbal instructions of acoustic
voice production measurement tasks. The researchers hypothesized that the amount of time and number of cues required to obtain acoustic measurements would be reduced and the stability of measurements would be increased in the experimental group due to prior video modeling.
CHAPTER II

Methods

Participants

Thirty-four young adult participants (mean age=21.32 years; SD=1.61) were enrolled in the study. Participants were recruited from Miami University, Oxford, Ohio. A randomized order was created by drawing from a bag containing an equal amount of papers labeled “Experimental” and “Control” to ensure equal group assignment for male and female participants. Upon enrollment in the study, participants were stratified based upon gender and then randomly assigned to either the experimental group or the control group. One participant was disqualified from participation due to hearing loss. Data from another participant was not used due to equipment malfunction during the course of the session. The analyzed data contained an equal distribution of male and female participants (8 male experimental, 8 male control, 8 female experimental, 8 female control).

Criteria for participation. Recruited participants were excluded from the study if they reported a history of smoking, current respiratory disease (for example, asthma, cystic fibrosis), or current neuromuscular disease (for example, multiple sclerosis) due to potential impact on voice production. Participants were also excluded from the study if they (1) currently or previously had been diagnosed with a voice or resonance disorder, (2) had been assessed for a voice disorder within the last five years, (3) had a hearing loss, or (4) had ever received a diagnosis of a learning disability. Individuals reporting an upper respiratory infection or exhibiting a perceptually dysphonic voice (as evaluated by the researcher) on the day of testing were excluded. Participants were also excluded if they reported a history of singing training or if they are a student majoring in speech pathology and audiology due to potential familiarity with the tasks included in the acoustic assessment protocol.

Procedure

Participants were recruited through the use of fliers posted in academic buildings in approved locations on the Miami University campus, informational emails dispersed through campus organization listservs, and word of mouth. Upon voicing interest in participating in the study, participants were sent an informational document via email. The document provided additional information regarding inclusionary/exclusionary criteria and the purpose of the study. Individuals who were interested in the study then contacted the Clinical Voice Laboratory at Miami University.
to schedule an appointment to complete the study. The research took place in the Clinical Voice Laboratory in the Department of Speech Pathology and Audiology, Bachelor Hall, Miami University.

All participants began in Room A with a research assistant. The research assistant completed all pre-experimental procedures with the participants in Room A. The researcher was not present in Room A during the informed consent, questionnaire, hearing screening, voice screening, and video training. The researcher was blind to the group assignment of the participant.

**Informed consent.** Each participant was given a written document of informed consent, explaining the details of the study including each task of the study protocol. The research assistant reviewed the consent document with the participant and provided ample time for the participant to read the document and ask questions. Participants were made aware through the consent form and verbal description of the study that participation was voluntary and that withdrawal from the study could occur at any time without penalty. Participants signed the informed consent document.

**Questionnaire.** Participants completed a short questionnaire (Appendix A) to validate inclusion in the study. Questions were related to singing experience, history of voice assessment and/or therapy, health and hearing status, and educational history.

**Hearing screening.** Prior to collecting the experimental sample, each participant was required to pass a bilateral pure-tone hearing screening. A hearing loss may cause inaccurate feedback for voice productions, thus skewing the data. A portable pure-tone audiometer (Maico Diagnostics, model MA25) was used to screen each participant. The participant was instructed to raise his or her hand whenever a tone was audible. Pure tones were presented at 25 dB HL and each participant was screened at 1000 Hz, 2000 Hz, and 4000 Hz. If the participant failed to raise his or her hand in response to one or more of the tones, the participant was excluded from the study.

**Voice screening.** Following the hearing screening, each participant was required to pass a voice screening for inclusion in the study. The Consensus Auditory-Perceptual Evaluation of Voice was used to screen participants for voice disorders. Perceptual qualities of voice, for example, roughness and breathiness, were examined. Prior to the completion of the study protocol, the participants read the first three lines of “The Rainbow Passage” into a microphone. The speech sample was recorded and a certified speech-language pathologist overseeing the study rated the sample if any discrepancy from normal voice was perceived.
Testing protocol: Control group. Following the voice screening, the research assistant escorted participants in the control group to Room B to begin the testing protocol.

Testing protocol: Experimental group. Following the voice screening, participants in the experimental group watched a training video on an iPad. The training video instructed the participants on completing acoustic voice assessment tasks that are used for obtaining fundamental frequency, intensity, maximum phonation time, and phonation range measurements. The video provided both verbal instructions and a video model of each task in the procedure. The purpose of watching the training video prior to beginning the acoustic voice assessment tasks was to simulate a technique that could be utilized during a real-life voice assessment. A patient undergoing a voice assessment could watch a training video on a tablet in the waiting room prior to their interaction with the clinician. Watching a training video would allow the patient to receive instruction for voice assessment tasks prior to beginning their session.

After the completion of the training video, the research assistant then escorted the experimental group participants to Room B.

Acoustic measurements. In Room B, all participants were seated in a quiet room containing the clinical voice assessment software (Real Time Pitch and Multidimensional Dimension Voice Program, KayPENTAX). Using a script, (Appendix B), the researcher provided all participants with verbal instructions and a model for each acoustic voice assessment task. The participants were given a practice trial for each task. The participants performed the task until they achieved success as judged by the researcher. Three trials of each task were performed. Tasks were instructed one at a time, so that the participants heard the instructions, watched a model, practiced, and then completed the task three times. The researcher held the microphone for the participant to achieve a consistent mouth-to microphone distance of 9 cm. These procedures mimicked traditional voice assessment procedures for obtaining acoustic voice measurements.

Five tasks were completed as part of the testing protocol. First, participants were asked to sustain an /a/ for approximately five seconds to measure fundamental frequency and intensity. Next, they read six short sentences (from the CAPE-V protocol) to assess speaking fundamental frequency and intensity. Participants were then asked to sustain an /a/ for as long as possible to evaluate maximum phonation time. Finally, participants were instructed to glide to their highest note on “whoop,” to determine their maximum fundamental frequency and glide to their lowest
note on “boom” to determine their minimum fundamental frequency. The session was video recorded on an iPad.

**Analysis of Data**

Several dependent variables were measured during data analysis. The first dependent measure was the amount of instructional time. Instructional time was defined as the session duration minus the time spent completing voice assessment tasks and extraneous interactions outside of a typical voice assessment (i.e. excessive talking, sneezing, etc.). To determine instructional time, each video recorded session was watched by the researcher. The total time of the session was recorded and each time any extraneous event occurred, the duration was recorded and subtracted from the total time. The time spent completing each task was derived from the recorded voice samples using Real-Time Pitch (KayPENTAX), which was also subtracted from the total time. A second dependent variable was cueing, defined as any verbal cues provided to the participant or questions asked by the participant and answered by the researcher. The number of cues was also recorded as the researcher watched the video recorded sessions. A primary goal of data analysis was to determine if the amount of time and number of cues required to successfully complete each task varied depending on the presence or absence of the training video. Inter-rater reliability for measurement of instructional time was achieved by having a second scorer (first year graduate student trained in measuring time) rate 20 percent of the samples. Seven videos were watched and analyzed by both scorers as training prior to the independent ratings by both scorers. Intra-rater reliability was achieved by having the primary rater score an additional 20 percent of samples a week after initial scoring. Independent t-tests were performed with reliability data to determine significance.

Other dependent measures included the mean values for fundamental frequency, intensity, maximum phonation time, maximum fundamental frequency, and minimum fundamental frequency as well as the average standard deviation for each measure. These measurements were derived from recorded voice samples using Real Time Pitch (KayPENTAX). The second goal of data analysis was to determine if the presence or absence of the training video impacted the average measurement and the stability of measurements across trials.

SPSS Version 21 was used for all statistical analyses. Independent samples t-tests were completed for each dependent variable. The alpha level for all statistical analyses was set at .05.
CHAPTER III

Results

Acoustic Assessment Completion Time

The mean instructional time values in the experimental versus control group participants are displayed in Figure 1. An independent t-test did not reveal a significant difference between the two groups (mean difference=.09 seconds, \( t=3.69 \), \( p=.714 \)).

Figure 1

*Mean Instructional Time*

![Bar chart showing mean instructional time for control and experimental groups.](chart.png)

Number of Cues

The mean number of cues required of the experimental versus control group participants are displayed in Figure 2. An independent t-test did not reveal a significant difference between the two groups (mean difference=-.69 cues, \( t=-1.02 \), \( p=.32 \)).
Figure 2

Mean Number of Cues

<table>
<thead>
<tr>
<th>Number of Cues</th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>2.31</td>
</tr>
</tbody>
</table>

Acoustic Measurement Means

Fundamental frequency. The mean fundamental frequency values in the experimental versus control group participants are displayed in Figure 3. An independent t-test did not reveal a significant difference between the two groups (mean difference = -.73 Hz, \( t = -0.4 \), \( p = .97 \)).

Figure 3

Mean Fundamental Frequency During Sustained /a/ Task

<table>
<thead>
<tr>
<th>Mean Hz</th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>162.41</td>
<td>161.67</td>
</tr>
</tbody>
</table>
**Intensity.** The mean intensity values in the experimental versus control group participants are displayed in Figure 4. An independent t-test did not reveal a significant difference between the two groups (mean difference=-2.50 dB, $t=-2.03$, $p=.051$).

Figure 4
*Mean Intensity During Sustained /a/ Task*

**Speaking fundamental frequency.** The mean speaking fundamental frequency values in the experimental versus control group participants are displayed in Figure 5. An independent t-test did not reveal a significant difference between the two groups (mean difference=3.91 Hz, $t=.22$, $p=.831$).
Figure 5

*Mean Speaking Fundamental Frequency*

![Mean Speaking Fundamental Frequency](image)

**Speaking intensity.** The mean speaking intensities in the experimental versus control group participants are displayed in Figure 6. An independent t-test did not reveal a significant difference between the two groups (mean difference=−.45 dB, t=−.44, p=.664).

Figure 6

*Mean Speaking Intensity*

![Mean Speaking Intensity](image)
**Maximum phonation time.** The mean maximum phonation times in the experimental versus control group participants are displayed in Figure 7. An independent t-test did not reveal a significant difference between the two groups (mean difference=-2.87 seconds, \( t=-1.31, p=.20 \)).

![Figure 7](image)

*Mean Maximum Phonation Time*

**Maximum frequency.** The mean maximum frequencies in the experimental versus control group participants are displayed in Figure 8. An independent t-test did not reveal a significant difference between the two groups (mean difference=21.79 Hz, \( t=.46, p=.65 \)).
Minimum frequency. The mean minimum frequency values in the experimental versus control group participants are displayed in Figure 9. An independent t-test did not reveal a significant difference between the two groups (mean difference=10.23 Hz, $t=.69$, $p=.498$).
Acoustic Measurement Standard Deviations

**Fundamental frequency.** The fundamental frequency standard deviation values in the experimental versus control group participants are displayed in Figure 10. An independent t-test did not reveal a significant difference between the two groups (mean difference=.99 SD, \( t=.99, p=.33 \)).

Figure 10

*Standard Deviation of Fundamental Frequency Measurements During Sustained /a/ Task*

![Bar graph showing standard deviation of fundamental frequency measurements for control and experimental groups.](image)

**Intensity.** The intensity standard deviation values in the experimental versus control group participants are displayed in Figure 11. An independent t-test did not reveal a significant difference between the two groups (mean difference=-.17 SD, \( t=-.76, p=.46 \)).
Speaking fundamental frequency. The speaking fundamental frequency standard deviation values in the experimental versus control group participants are displayed in Figure 12. An independent t-test did not reveal a significant difference between the two groups (mean difference=1.03 SD, \( t=1.26 \ p=.22 \)).
**Speaking Intensity.** The speaking intensity standard deviation values in the experimental versus control group participants are displayed in Figure 13. An independent t-test did not reveal a significant difference between the two groups (mean difference=.21 SD, \( t=-.99, p=.33 \)).

Figure 13

*Standard Deviation of Speaking Intensity*

![Bar chart showing standard deviation of speaking intensity for control and experimental groups.](image)

**Maximum phonation time.** The maximum phonation time standard deviation values in the experimental versus control group participants are displayed in Figure 14. An independent t-test did not reveal a significant difference between the two groups (mean difference=.65 SD, \( t=-1.26, p=.22 \)).
Figure 14

*Standard Deviation of Maximum Phonation Time Measurements*

Maximum frequency. The maximum frequency standard deviation values in the experimental versus control group participants are displayed in Figure 15. An independent t-test did not reveal a significant difference between the two groups (mean difference=−16.40 SD, t=−1.30, p=.20).

Figure 15

*Standard Deviation of Maximum Frequency Measurements*
**Minimum frequency.** The minimum frequency standard deviation values in the experimental versus control group participants are displayed in Figure 16. An independent t-test did not reveal a significant difference between the two groups (mean difference=-1.14 SD, \( t=-.83, p=.42 \)).

Figure 16

*Standard Deviation of Minimum Frequency Measurements*

<table>
<thead>
<tr>
<th>Average Standard Deviation</th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.98</td>
<td>3.84</td>
</tr>
</tbody>
</table>

**Reliability**

Inter-rater and intra-rater reliability were measured to determine the validity of study results. An independent t-test did not reveal a significant difference between instructional time measurements across two raters (mean difference=-.29, \( t=-.02, p=.98 \)). An independent t-test did not reveal a significant difference between instructional time measurements by the same raters on two dates at least one week apart (mean difference=3.14, \( t=.20, p=.85 \)).
CHAPTER IV

Discussion

The purpose of this study was to examine the impact of using video modeling techniques on session efficiency and stability of measurements during acoustic voice assessment. Data was obtained from 32 participants, ages 18 to 30 years, who were stratified based upon gender and then randomly assigned to the experimental or control groups. Participants in both groups began in Room A with the research assistant, where they gave informed consent, completed a questionnaire to confirm eligibility for study participation, and completed a hearing and voice screening. Participants in the experimental group then watched a training video that provided instruction and models for the completion of acoustic voice assessment tasks. Following the video, they proceeded to Room B. Participants in the control group immediately proceeded to Room B following the voice assessment. In Room B, participants in both groups received instruction, models, and cues from the researcher, who was blind to the participants’ group assignment, to complete acoustic voice assessment tasks. A variety of acoustic voice measurements were derived from recordings of three trials of each task. Instructional time and cues were derived from a video recording of the session.

Instructional time and cues

It was hypothesized that participants who watched a training video prior to completing acoustic voice assessment tasks would require less instructional time and less cues than those participants who did not watch a training video. Findings suggest that participants who had watched a training video did not require significantly less instructional time than their non-video-watching counterparts. Instructional time was calculated by taking the entire session duration and subtracting the amount of time spent completing tasks and extraneous talking/activities. Similarly, participants who watched a training video did not require significantly less cues. It is important to note that while the results were not significant, the mean number of cues required to successfully complete the acoustic voice assessment tasks was smaller for the experimental group.

Acoustic measurements

It was hypothesized that the acoustic measurements from participants who watched a training video prior to completing acoustic voice assessment tasks would be more stable than the acoustic measurements of participants who did not watch a training video. Stability was
determined by comparing the average standard deviation across three trials of the experimental group and control group. Mean and standard deviation measurements of fundamental frequency, intensity, maximum phonation time, maximum frequency, and minimum frequency were derived from digital recordings of acoustic voice assessment tasks. While none of the mean or standard deviation comparisons of the experimental and control group were found to be statistically significant, several measurements showed a trend towards more stability in the experimental group measures. The acoustic measurements trending towards more stability in the experimental group include mean intensity during the short, sustained /a/, mean maximum phonation time, standard deviation of fundamental frequency measurements during the sentence-reading task, standard deviation of maximum phonation time, and standard deviation of the maximum frequency. The tendencies suggest that while significant differences were not found between the video-watching and non-video-watching groups, mean values and stability of measurements may be areas to examine in similar tasks with other populations in future research.

Interesting to note is that three of the five measurements that exhibited a trend towards higher stability compared to other measurements in this study were found on maximum effort tasks. Maximum effort tasks, or maximum performance tests of speech production, are described as tests used to determine the upper boundaries that individuals are capable of producing with their voices (Kent et al., 1987). Maximum phonation time and phonation range (which includes maximum frequency) are considered to be maximum performance tests of speech production (Kent et al., 1987). It is with maximum effort tasks that a practice effect is often seen. Maximum effort tasks are notorious for large degrees of variability from trial to trial and across individuals (Kent et al., 1987). Thus, the difference in mean maximum phonation times between the two groups as well as the smaller standard deviations in maximum phonation time and maximum frequency measurements in the experimental group may be important to examine more closely. Classical literature within the field of speech-language pathology on the subject of maximum effort tasks cites inconsistent instructions to participants as well as a practice effect as possible causes of variability during maximum effort tasks (Kent et al., 1987). As such, it is important to consider the possibility that watching an instructional video prior to completing the maximum effort tasks may remove some intensity of the practice effect by providing more exposure to the task prior to completing each trial as well as providing more consistent instructions.
Limitations and Future Directions

Limitations of the present study are related to the breadth of the population studied. One area of concern with the present population is that the population was primarily comprised of college-students, presumably a group with much practice in learning novel tasks. Furthermore, the population studied presumably has consistent access to and experience with using digital media for the purpose of learning. As such, age and potentially situation (i.e., college students vs. non-college students) may be considered a limitation of the study. In future studies, it may be interesting to see if age differences, such as pediatric and older adult populations, have an effect on the time, cues, average acoustic measurements, and stability of measurements. While pediatric populations may have access to digital media in the school setting, young children inherently have less practice learning novel tasks compared to older populations. At the other end of the spectrum, an older adult population may have very different experiences with digital media for the purpose of learning.

Similarly, gender differences should also be examined in regard to learning outcomes with the use of a training video. The findings of this study would be enhanced through gender analysis to determine if significant differences exist between male and female experimental and control group participants. ANOVA tests should be used to examine gender effects in future analyses.

One final area that should be further examined in future studies is the effect of additional practices on task performance. In the present study, all participants were then given one practice following instruction of a task and a model. They then received cues and additional practice as necessary, judged by the researcher. Cues were taken into account in the analyses performed, but additional practices were not. In future studies, analysis of additional practice trials should occur.

Summary

This study examined the impact of using video modeling techniques on the session efficiency and stability of measurements during acoustic voice assessment. Thirty-four participants between the ages of 18 and 30 years were recruited for participation in the study and randomly assigned to either the treatment or control group for comparative analysis. Experimental group participants watched a training video prior to receiving verbal instruction and completing acoustic voice assessment tasks. Control group participants received the same
conditions with the exception of watching a training video. Findings from Independent Samples T-tests suggest that there is not a significant difference between the amount of time, number of cues, mean acoustic voice measurement values, and standard deviations across three trials when comparing those who have watched a training video and those who have not. However, a trend towards more stable results was seen on some tasks in the experimental group. Of particular interest is the trend seen on maximum effort tasks, such as maximum phonation time and maximum frequency. Different mean values for maximum phonation time and smaller average standard deviations across trials for maximum phonation time and maximum frequency were observed. These findings are important, because maximum effort tasks are notorious for variability between individuals and across trials for single individuals (Kent et al., 1987). Limitations in the present study include the homogeneity of the age studied and the potential gender differences that have not been examined. Directions for future study include expanding the age-range studied to include pediatric and older adult participants for a comparison across age groups to determine if there is a significant difference between the instructional time and consistency of measurements for children, young adults, and older adults. Gender analysis would also enhance the findings of this study. The use of a training video as part of a comprehensive voice assessment may have powerful implications to improve the efficiency and effectiveness of clinical voice assessment.
References


Appendix A
Questionnaire

Please answer the following questions to the best of your ability:

1. Gender (circle one): Male/Female
2. Date of Birth: ________________
3. Are you a smoker? (circle one): Yes/No
4. Have you ever been diagnosed with a voice disorder? (circle one): Yes/No
5. In the last 5 years, have you been assessed by a speech-language pathologist, ENT, or otolaryngologist regarding concerns about your voice? (circle one): Yes/No
6. Have you ever been diagnosed with any type or degree of hearing loss (circle one): Yes/No
7. Have you ever been diagnosed with a learning disability? (circle one): Yes/No
8. Have you ever been diagnosed with a neuromuscular disease? (circle one): Yes/No
9. Have you ever been diagnosed with a respiratory disorder? (circle one): Yes/No
10. Do you currently have any signs of an upper respiratory infection? (circle one): Yes/No
11. Have you ever received any sort of private vocal training? (circle one): Yes/No
12. Are you in the process of obtaining a degree in Speech Pathology and Audiology or have you ever received a degree in Speech Pathology and Audiology? (circle one): Yes/No
Appendix B

Script

Thank you for participating in our research study. Today, we will be doing three simple tasks to assess different aspects of your voice.

1: Fundamental Frequency/Intensity
The first task I will be asking you to complete allows us to find out about the pitch and loudness of your voice. I will hold the microphone next to your mouth. First we will have you hold out an /a/ for approximately 5 seconds, like this. (model). Now it’s your turn. I’ll have you practice and then we will complete the task. You will do the task three times.
(Practice and complete three trials)
Next, you will read 6 sentences. Try to speak at a comfortable level, similar to what you would use in your everyday conversations. Not too loud, soft, low, or high. I’ll model a few sentences for you. (model). Now it’s your turn. You practice and then we’ll do this task three times as well. Again, if you have any questions about the task, let me know.
(Practice and complete three trials)

2: Max Phonation Time
For our next task, you will again sustain an /a/ at a comfortable pitch for you. You will hold out this note as long as you can. Take a large breath in through your nose and really push from your abdominals until you can no longer produce any voice. We want to see the absolute longest amount of time you can sustain a note on one breath, so it is important that you make sure you take the largest breath possible and squeeze out all of the air to the point that you cannot push any more air out. This task will also be completed three times. I’ll model this for you (model).
(Practice and complete three trials)

3: Range: Real Time Pitch
For the next task, we will be assessing the highest and lowest notes you are able to make. Starting on a note that is comfortable for you, you will glide up the scale on a “whoop.” Make
your glide nice and smooth, without jumping from one note to the next. Really stretch it out, nice and smooth. I’ll model (model). We will repeat this three times. (Practice and complete three trials)

For the second part of the task, we will do the opposite. You will start on a note comfortable for you and glide down the scale on a “boom.” Like before, stretch out the glide and make the glide smooth, without lots of jumping from one note to the next. Make sure you don’t growl at the bottom, like this (gravelly noise). Rather end on your lowest note that is still comfortable to produce. I’ll demonstrate a correct glide down. (model) We will also do this task three times. (Practice and complete three times)

We’re all finished. Do you have any questions regarding the study? Thank you for your participation.