2005

Reliability Of The Concensus Auditory-perceptual Evaluation Of Voice On Pediatric Voices With Trained And Untrained Listeners

Teresa Elizabeth Pitts
University of Central Florida

RELIABILITY OF THE CONCENSUS AUDITORY-PERCEPTUAL EVALUATION OF VOICE ON PEDIATRIC VOICES WITH TRAINED AND UNTRAINED LISTENERS

by

TERESA E. PITTS
B.A. University of Central Florida, 2002

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Arts
in the Department of Communicative Disorders
in the College of Health and Public Affairs
at the University of Central Florida
Orlando, Florida

Summer Term
2005

ABSTRACT

Rating scales are commonly used to study voice quality. The purpose of this study was to examine inter-rater reliability/agreement of graduate student clinicians with differing levels of experience in rating voice perception. The Consensus Auditory Perception Evaluation of Voice (CAPE-V) was used to asses 1.) overall severity, 2.) roughness, 3.) breathiness, 4.) strain, 5.) pitch and 6.) loudness from a sample of pediatric voices. Twenty-four graduate clinicians who had completed a graduate level course in voice disorders participated in the study. Twelve of the participants were randomly selected to complete a perceptual training course prior to the evaluation session. Voice samples included 10 disordered and 2 normal voices from a population of children age 3-10 years old. The 12 voice samples were randomly repeated 3 times. Results of analysis of variance indicated that the groups significantly differed in their severity rating of the perceptual indices, suggesting that training affected the participants' judgment of severity. Additionally, variability was reduced as a function of training. Pearson product-moment correlation coefficient's revealed a moderate to strong relationship for all of the perceptual indices suggesting that regardless of training participants have an implicit understanding of normal versus disordered voice samples.

This work is dedicated to my dear friend

Sandy Rudasill.

In life and death you taught me to dream big.

You will *always* be remembered.

ACKNOWLEDGMENTS

I would like to thank Dr. Bari Hoffman-Ruddy at the University of Central Florida for your patience, guidance and encouragement; you have inspired me to continue to ask questions and strive for answers. Thanks to Dr. Jack Ryalls and Dr. Cecyle Perry-Carson for your time and recommendations. Dr. Celia Santini, you have been a mentor and a friend; thank you for everything. Dr. Alejandro Brice, thank you for all your help with the statistics. To my family, and Lindsay, thank you for the laughs and support that I needed to get through; especially Mom and Jessica who helped me conquer what seemed like an endless stream of numbers. To all the members of TORQ, I thank you for all your prayers and words of encouragement. Thank you all for believing in me. Additionally, I would like to thank all the University of Central Florida Communicative Disorders graduate students who participated in this study. Without you it would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

ASHA	American Speech and Hearing Association
BVA	British Voice Association
CAPE-V	Consensus Auditory-Perceptual Evaluation of Voice
MI	Mildly Deviant
MO	Moderately Deviant
SE	Severely Deviant
VPA	Vocal Profile Analysis

CHAPTER ONE: INTRODUCTION

In the field of speech pathology the human ear has been considered the 'Gold Standard' by clinicians and researchers for evaluating vocal quality. Historically, the evaluation and treatment of voice disorders have focused largely on the reliability of the clinician's perceived severity of the disorder rather than the use of instrumentation (Kreiman, Gerratt, Kempster, Erman, & Berke, 1993). More recently, the trend in the literature has been to utilize instrumental measures which relate to the clinician's perception of deviant voice quality. In this manner confirming the presence of a voice disorder is based on physical or physiological findings, rather than relying entirely upon an evaluator's perception.

Voice quality is a perceptual phenomenon that occurs in response to a given voice stimulus. Sound quality, in general, is defined as "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar" (ANSI Standard, 1960, p. 45). In a typical experiment on voice quality, listeners are presented with a set of stimuli and are asked to make a judgment of voice quality using some instrument such as a rating scale. The ratings made by the listeners are then used to interpret how they perceive voice quality.

In an effort to increase the reliability of the perceptual ratings determined by clinicians, various scales have been created to focus on and describe specific aspects of voice quality. To date, 57 scales have been used in the United States and the United Kingdom to evaluate voice disorders (Carding, Carlson, Epstein, Mathieson, & Shelwell, 2000). A large literature base has been developed to study the reliability of various perceptual evaluation scales and the three most

common scales reported on in the literature are the GRBAS scale, the Vocal Profile Analysis (VPA) and the BUFFALO-III Voice Profile.

In a study performed by the British Voice Association (BVA), these three scales were examined in order to develop a recommendation for speech-language pathologists in the United Kingdom concerning which of the scales should be utilized in their clinical practices (Carding et al., 2000). The authors' recommendations at the conclusion of the study indicated there was a strong need for the development of a new perceptual scale. This recommendation is based on the probability that the GRBAS, VPA or the BUFFALO-III do not fulfill all the needs of clinicians to adequately complete the task. Furthermore, the authors concluded that a rating system needed to be developed which is clinically realistic, theoretically sound, and proven reliable (Carding et al., 2000).

In 2002, The American Speech-Language-Hearing Association's special interest division (Division 3) of voice and voice disorders and experts in human perception conferred to develop guidelines for standardizing the perceptual evaluation of voice disorders. The outcome of this meeting was the creation of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), with the ultimate goal of being a reliable tool for use in clinical practice. To date, there are no published reports which have evaluated the reliability or validity of the CAPE-V with pediatric voices.

Reliability of Perceptual Measurements

In order to evaluate the reliability of the perceptual indices included in the CAPE-V, it is necessary to first examine previous research conducted on the reliability of perceptual measurements. Blaustein and Bar, (1983) tested the reliability of perceptual voice assessments

using experienced clinicians and 161 school-age children who had been screened for phonatory aspects which were considered to be a voice disorder. The study did not define the severity ratings to allow clinicians use of their professional internal standards to specifically examine intra/ inter judge reliability of speech pathologists. This study concluded that inter-judge agreement varied for each measurement. The inter-judge agreement in the areas of vocal quality, breathiness and nasal resonance was low, there was reasonable agreement for hoarseness and stridency, and finally, there was significant agreement in changes of pitch. The authors of the study concluded the research displayed a "wide variation" between the various clinical judgments and called for the creation of an objective method for the assessment of disordered voices (Blaustein & Bar, 1983).

Kreiman and colleagues (1993) examined the use internal standards to anchor a clinician's judgment. Disordered voices were rated on perceptual scales by two groups of participants. Group one was given anchors before the sessions began. The anchors were examples of the five points of severity along the continuum of the rating scale. When the anchors were used as a basis for evaluation little drift was noted during the session, as opposed to substantial drifts noted in the opposing group which was not provided anchors. Based on these results, the authors concluded, to increase reliability anchors or explicitly defined parameters should be provided. Blaustein and Bar (1983) and Kreiman et al. (1993) discussed investigating the correlation between a clinician's level of experience and the reliability of their conclusions. The authors stated that there was a need for clinicians to develop internal standards as a method to anchor their perceptual judgments.

Further research has been conducted to study the correlation between the level of experience/ training and the reliability of a clinician's judgment. In a study conducted by

DeBoer and Shealy (1995), ten speech-language pathology graduate students were given seven weeks of singing lessons. The purpose of the study was to determine whether singing lessons, similar to those received from a private teacher, could have an affect upon a clinician's perceptual skills. The results of this study revealed that singing lessons had a significant impact on the clinician's ability to perceive the areas of pitch, breath control, legato production and easy onset. The significance was seen in the quantitative testing, subjective testing, student journal comments, and student evaluations. The subjective test results further revealed improvement in the ability to discriminate between different voice qualities. The study supported the idea of integrating singing lessons into graduate study to increase beginning clinician's ability to accurately judge specific aspects of voice quality.

Additionally, a study performed by Wolf, Martin, and Palmer (2000) examined varying perceptual dimensions perceived by native listeners. Native listeners included adults who had no formal training in voice or voice disorders. The participants listened to 20 dysphonic vowel samples. To investigate the perceptual differences of various attributes in the vocal sample, the researchers used the Multi-Dimensional Scaling Approach. This requires the participants to only judge the similarity or dissimilarity between voices, instead of giving severity ratings. The authors concluded the perceptual judgments of native listeners are affected more by the severity of the abnormality vs. the specific attribute being judged. The results also revealed five dimensions present in the voice sample accounted for 45% of the perceptual judgment variation noted: 1) hoarseness; 2) elevation of second formant; 3) breathiness/ nasality; 4) uninterpretability of the sample; and 5) fundamental frequency. When compared to the study performed by Kreiman, Gerratt, and Precoda (1990), which examined the results of the perceptual judgments from the expert listeners and the acoustic correlates, the acoustic quality

accounting for almost 50% of the judgment variation was fundamental frequency. In a similar study by Wolfe et al. (2000) the acoustic qualities accounting for the judgment variations were fundamental frequency, intensity, and perturbation. The authors concluded that the presence of a variation in fundamental frequency and/or intensity, other dimensions of voice quality have the potential to be masked.

Differences in Terminology

The congruency of terms used to define and describe differing aspects of the vocal sample is another important aspect to consider. In a study conducted by van der Torn, Verdonck-de Leeuw, Kuik, and Mahieu (2002) on communication suitability of voice following radiotherapy for T1 glottic carcinoma, the following subjective perceptual terms were used: "unpleasant/ pleasant"; "ugly/ beautiful"; "breathy/ not breathy"; "dull/ clear"; "high/ low"; "shrill/ deep"; "unsteady/ steady"; "panting/ tense"; "relaxed/ rough"; and "creaky/ deviant." Other terms that have been used are froggy, stridency, glottal, harsh, and hoarse.

A study by Bassich, and Ludlow (1986) investigated the level of training required for beginning clinicians to reliably rate specific dimensions of vocal quality. Graduate students were trained in the assessment the voice quality of the vowels /a/ and /i/ in normal and disordered voices, using 13 different dimensions of quality: 1) tremor waver, 2) rough-fry, 3) wet-horse, 4) harsh-shrill, 5) breathy, 6) strain-strangle, 7) pitch level, 8) loudness level, 9) pitch breaks, 10) nasality, 11) pitch stability, and 12) loudness stability, on a five point scale. Definitions and two normal voice and eight deviant voices for each dimension were presented to the participants. Sixteen one-half hour training sessions were required before the mean inter-judge reliability attained 80% for all dimensions. After the training sessions, 25 experimental sessions were

conducted to rate new normal and pathologic voices. The dimensions of tremor, loudness, breathy and wet-horse were seen to have reliability greater than 80% in the early training sessions, others such as harsh-shrill, rough-fry, and nasality met the requirement later in training. Only three dimensions: 1) strain-strangle, 2) loudness stability, and 3) loudness level, had inter-judge reliability coefficients greater than .90 for pathologic samples during the experimental session. The authors concluded the use of extensive training was required for the clinicians to achieve the 80% marker, and even when this marker was achieved their judgments did not remain stable. Additionally, the reliability of perceptual judgments seen could have been based on the minimal level of distinctness between a portion the chosen categories, for example, loudness stability and loudness level. They compared the present results with experienced clinicians and concluded perceptual rating of dysphonic voices is a difficult task and to increase reliability extensive professional experience and training is needed (Bassich, & Ludlow, 1986).

CAPE-V Perceptual Indices

The Consensus Auditory- Perceptual Evaluation of Voice (CAPE-V) uses the following perceptual terms: 1) roughness, 2) breathiness, 3) strain, 4) pitch, and 5) loudness. Roughness, which is also described in the literature as harshness, is the first perceptual term used in the CAPE-V. Roughness/ harshness can be defined in two areas: rough-fry which is a rough voice quality during low pitched phonations, which may be associated with a beating or cracking of glottal fry, or harsh-shrill. Additionally, it is described as a rough or strident voice quality during high pitched phonations, which can be associated with hard glottal attacks (Bassich, & Ludlow, 1986). The CAPE-V defines roughness simply as the presence of irregularity in the voice (ASHA Division 3, 2004).

A study attempting to predict vocal severity found the psychological cause of harshness is associated with turbulent airflow at the area of the glottis. The authors stated the cause could be found in several types of glottal area function, and the production of this quality is more complex than a breathy voice quality (Wolfe & Steinfatt, 1987). Sapienza, Hoffman-Ruddy, & Baker (2004) stated one cause of harshness in the vocal sample is the presence of vocal nodules. More than one million school age children have been diagnosed with vocal nodules, and in the pediatric patient they are typically the same color as the true vocal folds and appear on the on the free margin of the vocal fold at the junction of the anterior and middle third (Sapienza, Hoffman-Ruddy, & Baker; 2004).

Breathiness is described as an audible escape of air relating to the inability of the vocal folds to fully adduct resulting in a thin, weak phonation (Bassich & Ludlow, 1986). The CAPE-V defines breathiness as an audible escape of air in the voice (ASHA Division 3, 2004). In a study by Wolfe & Steinfatt (1987), attempting to predict the vocal severity across voice types, the authors discovered breathy voice characteristics were found in the following voice disorders: nodules, polyps, recurrent laryngeal nerve damage, edema, postpolypectomy dysphonia (no lesion), polypoid laryngitis plus nodule, or a contact granuloma.

Strain is described as phonation with effortful squeezing of the voice (Bassich & Ludlow, 1986). The CAPE-V describes strain as excessive vocal effort or hyperfunction (ASHA Division 3, 2004). Sapienza et al. (2004) stated laryngeal hyperfunction in children is caused by an increased tension in the extrinsic laryngeal musculature, with adduction of the false vocal folds and compression of the arytenoid cartilages. Overtime, hyperfunction can cause swelling or edema in the true vocal folds. In newborns and young children, the vocal fold mucosa is thinner,

and therefore, not a clear differentiation between the superficial, intermediate and deep layers of the lamina propria as compared to adults (Sapienza et al., 2004).

Pitch is defined as the level of phonation as compared to a vowel phonation for an individual's age and sex (Bassich & Ludlow, 1986). The perception of pitch is related to the size of the laryngeal structures. The length of newborns vocal folds are on average 2.5-3.0 mm and continually grow until around twenty years of age. Around ten to fourteen years, there is the appearance of sex differences in vocal fold length (Sapienza et al., 2004). The average fundamental frequency for an adult male is 126 Hz, with a standard deviation of 26.52 Hz. For the adult female, the average is 181 Hz with a standard deviation of 45.18 Hz (Xue & Fucci; 2000). For a child, the average fundamental frequency observed is 250 Hz (Zajac, Farkas, Dindzans, & Stool; 1993). Hufnagle (1982) found the fundamental frequency of normal children, between the ages of ages of six and eight, was 246.4 Hz with a standard deviation of 26.82 Hz, and when he examined children with a diagnoses of vocal nodules were matched for age and weight he found the mean was 301.44 Hz with a standard deviation of 49.62 Hz. The children with vocal nodules presented with significantly higher fundamental frequencies.

The final perceptual term included in the CAPE-V is loudness. The CAPE-V defines loudness as the perceptual correlate of sound intensity. The severity level of the loudness scale indicates whether there is deviation from normal based upon the patient's gender, age and referent culture (ASHA Division 3, 2004).

Perceptual Rating Scales

Prior to the introduction of the CAPE-V the most popular evaluation methods utilized in a clinical practice were the Vocal Profile Analysis (VPA), Buffalo III Voice Profile, and the

GRBAS scale (Carding et al., 2000). These scales vary considerably in terminology and the aspects of the voice samples being assessed.

The VPA was developed by a phonetician and a speech-language pathologist for use in analyzing voices in sentences and conversation (Carding et al., 2000). The definition of voice quality was presented in this scale was broader than had been traditionally used (Wirz, 1995). This broader definition of voice quality allowed for the assessment of the entire vocal apparatus including the laryngeal, supra-laryngeal, prosodic aspects, and how these different aspects may 'color' the perception of an individual's voice (Wirz, 1995).

The VPA required the clinician compare the voice of interest to a defined standard, and then give a rating for each parameter. The basis of this approach allowed for an objective basis of comparison (Wirz, 1995; Carding et al., 2000). Additionally, the VPA described settings, or independent components which may be combined in different ways to explain the overall impression of a voice (Wirz, 1995). The positive aspects of the VPA are its suitability for both normal and abnormal voices, and possible utilization by singing and voice teachers. However, it required a two day training program in order to train the clinician to effectively use the scale. In a study discussed by the British Voice Association which examined the reliability of the VPA, results demonstrated there was evidence of intra/inter judge reliability; however perceptual judges were required to attend regular training to maintain their level of reliability (Carding et al., 2000).

The Buffalo III Voice Profile was created for the specific evaluation of pediatric voices (McAllister, 1997). The scale examines twelve aspects of the voice quality: 1) laryngeal tone, 2) loudness, 3) nasal resonance, 4) oral resonance, 5) pitch, 6) breath supply, 7) muscles, 8) voice abuse, 9) rate, 10) speech anxiety, 11) speech intelligibility and 12) overall voice rating. These

twelve aspects are evaluated on a scale from one (normal), to five (very severe). The scale places emphasis on both the general aspects of vocal function and behavior (Carding et al., 2000).

The positive aspects of the Buffalo-III were its simplicity/ease of use and its broad range of categories. Negative statements of the scale were the inclusion of non-voice parameters, formalized training is not recommended, it is not applicable to normal voices, there were no audio tapes for listener training, the scale did not address the supra-laryngeal tract, and it was not applicable for singing and voice teachers (Carding et al., 2000).

The final scale to be considered here is the GRBAS scale. This scale was developed in an effort to explain the psychoacoustic phenomenon of hoarseness utilizing the Osgood Semantic Differential Technique (Wirz, 1995; Hirano, 1981). This work was developed further by Isshiki and the Japanese Society of Logopedics and Phoniatrics resulting in the GRBAS scale.

This scale evaluated five aspects of vocal quality listed below:

G-grade: “Degree of Abnormality

R-rough: "Irregularity of Fold Vibration"

B-breathy: "Air Leakage in the Glottis"

A-aesthinic: "Lack of Power"

S-strained: "Hyper Functional State" (Hirano, 1981).

For each of the five vocal parameters listed there is a four point severity scale ranging from zero to three provided for the clinician to make a designation regarding the severity of each feature (De Bodt, Wuyts, Heyning, & Croux, 1997). The British Voice Association discussed some of the positive aspects of the GRBAS scale including: its simplicity/ ease of use and well defined

terminology. Additionally, there is evidence of intra/inter judgment reliability; the perceptual terms chosen were pertinent to the laryngeal features, and the ability to rate degrees of abnormalities. Finally, the audio tapes provided for the training of clinicians were helpful and the scale's administration time is short, approximately five minutes (Carding et al., 2000). De Bodt et al. (1997) stated the scale was a compact perceptual rating system which can be used daily by all members of the voice care team (De Bodt et al., 1997).

Review of the literature, by the British Voice Association, indicated the GRBAS scale was limited to evaluating only at the laryngeal level, there is no component for ratings of the vocal tract specifically the supra-glottic parameters, and common features such as pitch and loudness were not included. Moreover, there is no formalized training in using the system, no formal protocol present, and it is not applicable to a voice or singing teacher (Carding et al., 2000).

De Bodt et al. (1997) retested the reliability of the GRBAS scale while examining the influence of professional background and experience. Nine pathological voice samples were presented to 23 judges. The judges consisted of experienced and inexperienced otolaryngologists and speech-pathologists. The sets of voices were presented with a two-week lapse or greater. The study demonstrated only moderate overall reliability; profession and experience did not influence the overall outcome; and breaking the results down further between the two specialties (speech-pathologists and otolaryngologists) did not significantly alter the results, but the level of experience among the groups did increase reliability above the moderate range, to what the authors stated was, good.

With regards to the specific terms, the results of this study demonstrated the overall rating for the severity of hoarseness was good and was the most consistent of all of the

parameters. In evaluating the parameters roughness and breathiness, the speech-language pathologists were found to be more reliable than otolaryngologists and their reliability increased with the number of years of experience. The reliability of the parameter aesthenia was less reliable than all other parameters. The authors of the study suggested aesthenia was the one area highly susceptible to listener error. The parameter strain also had low reliability, and its reliability did not appear to increase with experience (De Bodt et al., 1997).

Since the groups did not differ significantly, the authors suggested professionals with varying specialties and the level of experience can use this protocol to assist in the diagnosis of voice disorders. The authors also suggested because of the inconsistent internal standards, and the instability present in the ratings, clinicians required an average of eight hours of training to achieve 80% inter-judge reliability (De Bodt et al., 1997). In addition, Carding et al. (2000) and Wein (1981) stated the GRBAS scale should be the absolute minimum for perceptual evaluation of voice practiced in the United Kingdom, but improvements in the perceptual ratings were necessary.

The CAPE-V was developed by the American Speech-Language Association's (ASHA) Division 3: Voice and Voice Disorders (ASHA Division3, 2004). It was developed with the purpose of creating standardized guidelines for the auditory-perception of voice allowing for practicality and a short administration time. Design considerations included: perceptual dimensions should reflect a minimal set of clinically meaningful, perceptual voice parameters, identified by a group of expert clinicians. Secondly, the procedures and results should be obtainable expediently. Thirdly, procedures and results should be applicable to a broad range of vocal pathologies and clinical settings. Lastly, the ratings ultimately should be demonstrated to

optimize reliability within and across clinicians, and exemplars should be available for training (ASHA Division 3, 2004).

The CAPE-V examines six vocal attributes: 1) overall severity, 2) roughness, 3) breathiness, 4) strain, 5) pitch, and 6) loudness. The clinician denotes the level of severity by placing a mark along a 100 millimeter line using a visual analog scale. The mark further indicates a general region of severity consisting of "MI" or mildly deviant, "MO" or moderately deviant, and "SE" or severely deviant. To date published studies do not exist on the reliability of this perceptual scale in the adult and pediatric population. One study is currently in progress utilizing adult voices as the perceptual samples.

Purpose of this Study

The primary focus of this study is to examine the reliability of clinician's rating of perceptual parameters utilizing the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) protocol on a sample of pediatric voices. This study included participants with differing levels of experience in the field of speech pathology. Two groups were included: graduate students who have concluded a class in voice disorders; and graduate students who have concluded a class in voice disorders and undergone a training program for the perceptual judgment of voice disorders.

Hypotheses

This study tests the central hypothesis that training will increase listener agreement and reliability of perceived voice quality. Specifically:

1.) Ho: The untrained listeners will not be in agreement/ reliable when rating the perceptual indices, this will be tested with the Pearson product-moment correlation coefficient; r will be accepted as less than or equal to .10, a strong correlation will be observed from .80- 1.0.

Ha: The untrained listeners will be in agreement/ reliable when rating the perceptual indices, this will be tested with the Pearson product-moment correlation coefficient; r will be accepted at greater than or equal to .10, a strong correlation will be observed from .80- 1.0.

2.) Ho: The listeners will not be in agreement/ reliable when rating the perceptual indices, this will be tested with the Pearson product-moment correlation coefficient; r will be accepted as less than or equal to .10, a strong correlation will be observed from .80- 1.0.

Ha: The listeners will not be in agreement/ reliable when rating the perceptual indices, this will be tested with the Pearson product-moment correlation coefficient; r will be accepted at greater than or equal to .10, a strong correlation will be observed from .80- 1.0.

3.) Ho: The perceptual ratings of the trained group will not be significantly different as compared the untrained group; an Analysis of Variance will be used to test the difference between group 1 and group 2, and significance will be accepted at $p > .05$.

Ha: The perceptual ratings of the trained group will be significantly different as compared the untrained group; an Analysis of Variance will be used to test the difference between group 1 and group 2, and significance will be accepted at $p < .05$.

CHAPTER TWO: METHODOLOGY

Study Design

This study represents a prospective, repeated measure, randomized design with six factors and three levels. All participants made ratings on randomized voice samples from a pediatric population using the CAPE-V to study the reliability of this perceptual rating protocol. The independent variables were the perceptual terms included in the CAPE-V: 1) overall severity, 2) roughness, 3) breathiness, 4) strain, 5) pitch, and 6) loudness. The dependent variables included are the participant's levels of experience.

Participants

Twenty-four participants (male and female) age 18-35 were included in this study. The participants were divided into two groups. The 24 participants were selected from the student population at the University of Central Florida's Graduate program in the Department of Communicative Disorders located in Orlando, Florida. Participants were excluded from this study if a hearing or visual impairment was present.

The participants were randomly assigned to one of two groups. Group 1 was the untrained participants and group 2 underwent training of the perceptual evaluation of voice and the CAPE-V protocol.

Voice Samples

Voice samples were obtained from 10 children who were referred from the Seminole County Public School System for a voice evaluation. The children were seen at the University of Central Florida's Voice Care Center for a medical and laryngeal evaluation. Informed Consent was obtained from the guardian of each student before participating in the evaluation process. The child was asked to read the "Rainbow Passage" (Fairbanks, 1960), CAPE-V sentences, and sustain the vowel /a/. A cardoid headset microphone was used to collect the audio samples. The microphone was connected to a tube MP preamplifier and SONY TCD-D8 portable DAT recorder. Mouth-to-microphone distance was held constant for all participants. Distance was set at 1cm from the lips and placed at a 70 degree angle to the center of the lips. The children were asked to repeat this protocol three times. From these recordings 10 voice samples were selected with the following requirements: the students were required to be between the ages of 5 to 12 years old with a diagnosis made regarding the nature of the voice disorder (Appendix D). Two normal voices were also included in the study (n=12).

The voice samples were digitized at 48 mega-hertz using Sound Trac Software, and final edits were completed using Final Cut Pro version 3.1 (Apple Computer Inc., 2005). The 12 samples were assigned a number between 1 and 12. The voice samples were randomized for the second and third presentations using SPSS to decrease the learning effect from repetitive presentation orders (SPSS Inc, 2003). Additionally, the set presentation order of the entire 36 repeating samples was not changed across the participants, to allow for comparison within listeners (Kreiman, et al., 1993).

Procedures

The procedure for the perceptual training program included two groups. Group 2 (participants who received perceptual training) attended two training sessions at the University of Central Florida's Department of Communication Disorders, Vocal Function Laboratory. During the training sessions the participants were presented with information on the history of the CAPE-V and the definition and meaning for each of the perceptual terms. Additionally, the students listened to examples of disordered voices for each perceptual term and practiced making perceptual ratings. Feedback was provided from the instructor indicating the severity of the CAPE-V parameters. Following the second training session, the protocol of the CAPE-V was presented, along with instructions on how to complete the CAPE-V form, and then the students completed the experimental evaluations. For group 1 (the participants not receiving the training program) the same procedure was followed with the exception of the two training sessions.

For both groups, a monitor was present for all evaluation sessions conducted. The participants were seated in designated areas and given Sennheiser eH 1430 headphones and the samples were presented at 75dB. Directions for the session were given verbally, regarding the time allowed for the evaluation, use of equipment, and the directions on the CAPE-V protocol. The 0 mm mark represented a normal voice sample in the CAPE-V; the level of severity increased with increments along the scale. All areas on the CAPE-V form were marked for each voice sample.

The evaluation session lasted one hour, to complete the entire experimental protocol. The participants were allowed to listen to each vocal sample only once in entirety, and questions

were allowed to be asked during the evaluation time. The CAPE-V protocols were collected after each evaluation session.

Analysis and Scoring of Perceptual Data

After all the evaluations were completed on the visual analog scale, the ratings were calculated by the experimenter. The distance was physically measured in millimeters, from the far left end of the scale to the point where the clinician made the mark. The length of the line was written in the blank space provided. The line lengths were entered into data sets for statistical analysis.

CHAPTER THREE: RESULTS

Descriptive Statistics

The independent variable for this study was the participant's levels of experience in judging perceptual measures (group 1: untrained participants, group 2: trained participants). The dependent variables were the perceptual indices included in the CAPE-V: 1) overall severity, 2) roughness, 3) breathiness, 4) strain, 5) pitch, and 6) loudness. Descriptive statistics included mean and standard deviations for the dependent variables. Mean severity scores were also calculated for each group by combining the scores for each repetition, per perceptual indice, per group. There were a total of 144 samples, per group, for each of the six perceptual indices. Standard deviation was derived from 12 perceptual measures per dependent variable, for each of the repetitions of each voice sample. The results can be seen in Table 1-24. (The disordered voice samples are numbered 1-10, and the normal samples are numbers 11 and 12.)

The mean severity scores for overall severity (Table 1 & 2), from group 1 and 2, over the three repetitions, remained within a single severity range for each of the voice samples. For example, for group 1, sample 1, all participants' ratings when averaged calculated to be within 17-25mm, which is within the mild-moderate range on the CAPE-V visual analog scale. For group 2, sample 1, all participants' rating were calculated to be within 17-21mm which is within the mild-moderate range. Additionally, the voice samples that were rated in the moderate range (35-70mm), voice samples 2, 4, 5, 6, and 8, were marked 10-20mm higher by group 2 in comparison with group 1.

The mean severity scores for roughness (Table 5 & 6) are similar to what was noted for overall severity. There was an overall decrease in the ranking of severity for all the voice samples by group 2. The only exception was samples 11 and 12, which were the normal voice samples. The mean severity scores for breathiness (Table 9 & 10) presented in a similar manner when compared to overall severity and roughness. Voice samples that had severity scores in the moderate range, the samples were observed to be rated 10-20mm higher by group 1. This trend was also seen for the mean severity score for strain (Table 13 & 14) The mean severity scores for group 2 were on average 10-20mm lower than group 1, with the exception of voice sample 7.

Group 1 and group 2 were consistent in their mean severity ratings of pitch (Table 17 & 18); on average the mean severity difference over the three repetitions was 5mm for group 1 and 2mm for group 2. However, the mean severity scores for pitch were greatly reduced for group 1 and 2, in comparison with the results for the previous perceptual indices.

The mean severity scores for loudness (Table 21 & 22) were similar to what was noted for pitch. There was little difference between the mean severity scores over the three repetitions for group 1 and group 2. Interestingly, group 2 did not give a severity score over 13mm, which is a mild rating on the CAPE-V. Group 2 did not give a severity rating over 43mm which corresponds to an early moderate rating on the CAPE-V.

Variability

The inherent variability within human subjects for making perceptual ratings has been discussed in the open literature. Standard deviations were derived from the mean severity scores for group 1 and group 2, for each of the perceptual indices, for each of the three repetitions. Tables 3-4, 7-8, 11-12, 15-16, 19-20, and 23-24 depict standard deviation data for group by

perceptual indice. Overall, the standard deviations for both groups for the dependant variables were high. This demonstrates the variability of the participants in making perceptual judgments. For example, the standard deviations for roughness (Table 7 & 8) in group one was consistently within 15mm- 25mm for the disordered voice samples (1-10). With group 2 there was more variation observed. Many times the standard deviations for group 2 were substantially lower than group 1, (e.g., voice sample 1). In other instances, (e.g., voice sample 7), there was increased variability for the second and third repetition in group 2.

The standard deviations for pitch (Table 19 & 20) in group 2 were consistently lower than group 1, with the exception of voice sample 7 and the first repetition for voice sample 12. There was stronger agreement among group 2 that pitch variability was not present in the majority of voice samples presented. The standard deviations for loudness (Table 23 & 24) were observed to be similar when compared to pitch. Group 2 was considerably lower in comparison to group 1. For example, voice sample 2 fell was rated one standard deviation over the three repetitions (i.e., group 1: 24, 15 & 21; group 2: 10, 9 & 7).

Scatter plots were formulated to provide visual comparison of the mean to standard deviation data for each group including all perceptual indices (group 1: figure 1; group 2: figure 2). Overall, these plots indicated that group 2 rated the voice samples with decreased severity, and the majority of their standard deviations were less than 10 mm, when compared to group 1. The voice samples rated for group 1 appear to cluster in 2 categories: a cluster in a mild rating and a cluster in the moderate-severe rating, with the majority of the standard deviations less than 15-20mm.

Inferential Statistics

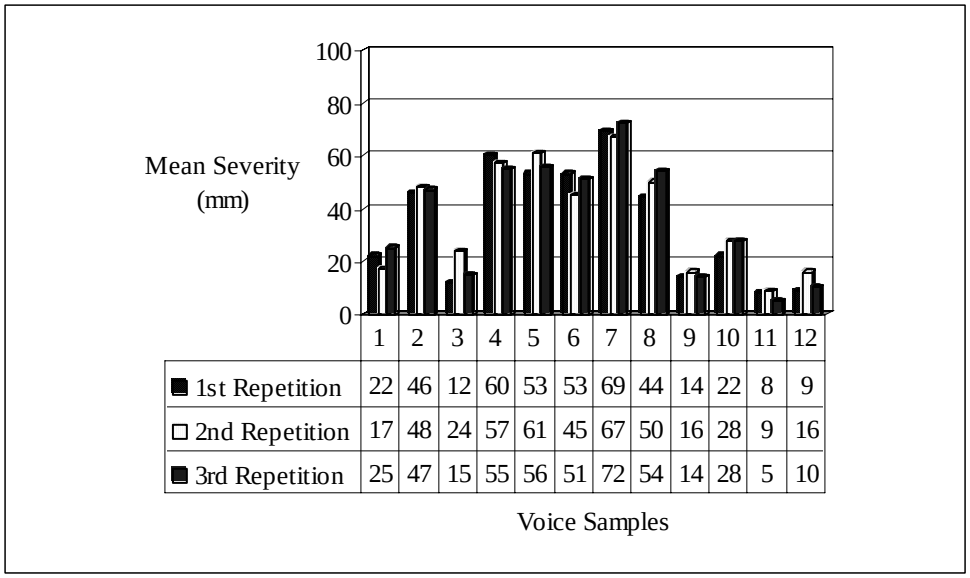
One-way Analysis of Variance (ANOVA) was used to test for significance of mean differences across all dependant variables. The data were analyzed using SPSS version 12.0 for Windows (SPSS Inc., 2003). ANOVA indicated that there were differences between group 1 and group 2 for several dependent variables. Results can be seen in table 25-30.

For overall severity (table 25) the voice samples 2, 4, 6, and 8 were statistically significant. For roughness (table 26) the voice samples 1, 4, and 6, were statistically significant. For breathiness (table 27) voice sample 3 was statistically significant. For strain (table 28) the voice samples 1, 2, 3, 4, 5, 6, 8 and 10, were statistically significant. For pitch (table 29) the voice samples 2, 4, 5, 6, 7, 8 and 12, were statistically significant. For loudness (table 30) the voice samples 1, 2, 4, 5, 6, 7, 8 and 12, were statistically significant.

Pearson product-moment correlation coefficients were run test inter-rater reliability for each group and to measure the strength of the correlation between the three repetitions of each voice sample (Kreiman, Gerratt, Kempster, Erman & Berke 1993). The magnitude of correlation was interpreted as follows: r values from 0.0 to 0.2 represent a very weak to negligible correlation; 0.2 to 0.4 represents a weak, low correlation; 0.5 to 0.7 represents a moderate correlation; 0.8 to 0.9 represents a strong, high correlation; and 0.9 to 1.0 represents a very strong correlation (Doehring, 1996; Maxwell & Satake, 1997; Schiavetti & Metz, 2002). A Pearson r correlation matrix depicts the results of the dependent variables, see Table 31 and figures 3-38. The correlations revealed that for 14 out of the 18 points of comparison, the correlations for group 2 were stronger than group1.

Table 1

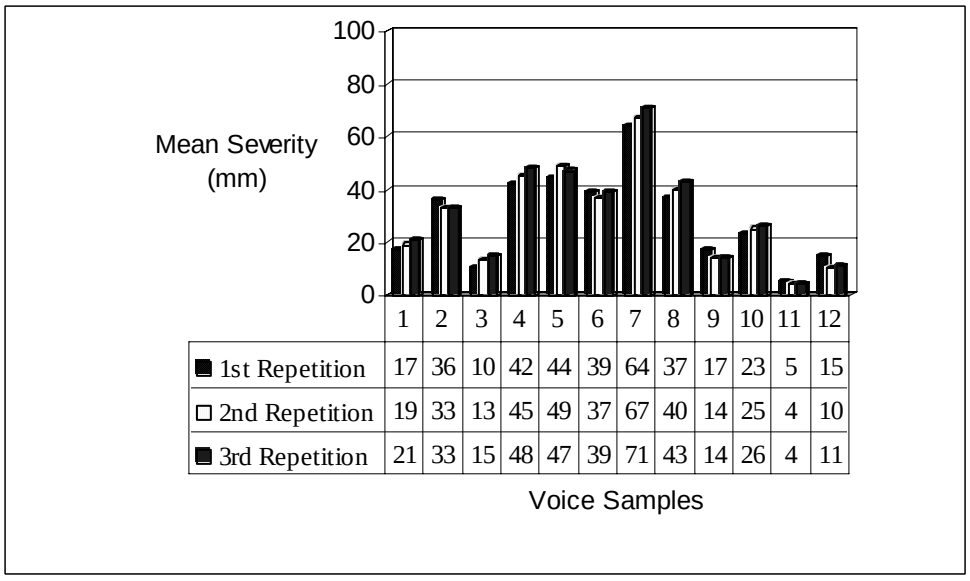
Mean Severity in Millimeters for Group 1 Rating of Overall Severity



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	22	46	12	60	53	53	69	44	14	22	8	9
2nd Repetition	17	48	24	57	61	45	67	50	16	28	9	16
3rd Repetition	25	47	15	55	56	51	72	54	14	28	5	10

Table 2

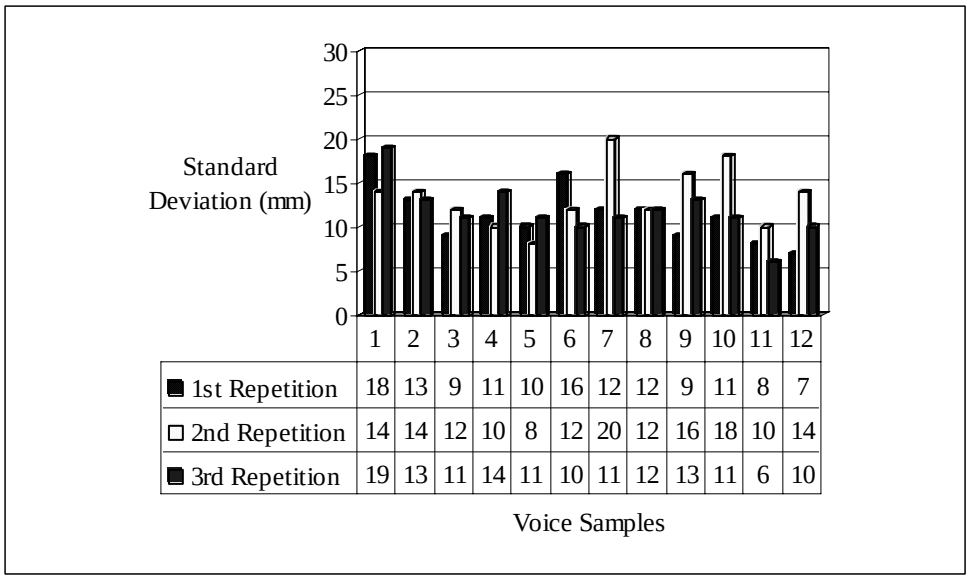
Mean Severity in Millimeters for Group 2 Rating of Overall Severity



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	17	36	10	42	44	39	64	37	17	23	5	15
2nd Repetition	19	33	13	45	49	37	67	40	14	25	4	10
3rd Repetition	21	33	15	48	47	39	71	43	14	26	4	11

Table 3

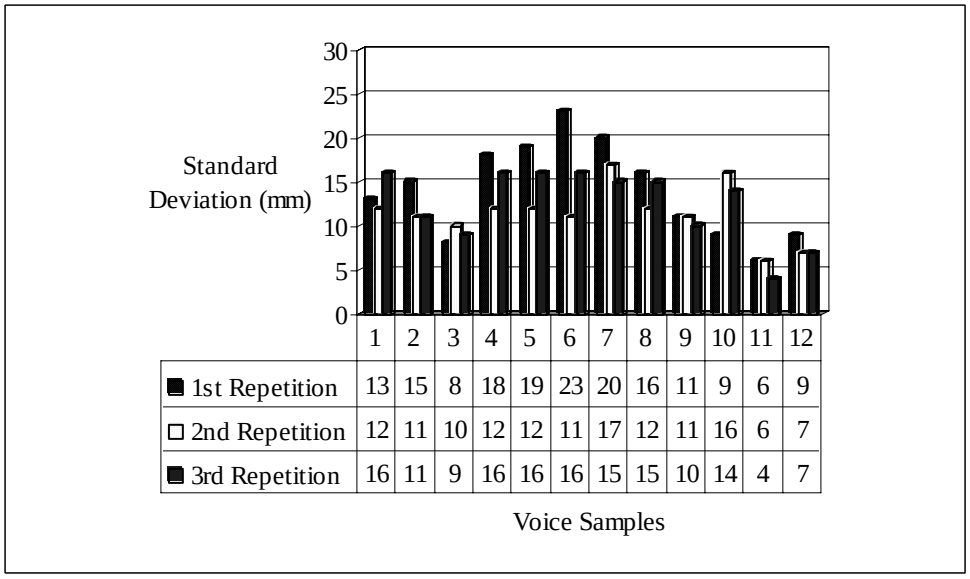
One Standard Deviation in Millimeters for Group 1 Rating of Overall Severity



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	18	13	9	11	10	16	12	12	9	11	8	7
2nd Repetition	14	14	12	10	8	12	20	12	16	18	10	14
3rd Repetition	19	13	11	14	11	10	11	12	13	11	6	10

Table 4

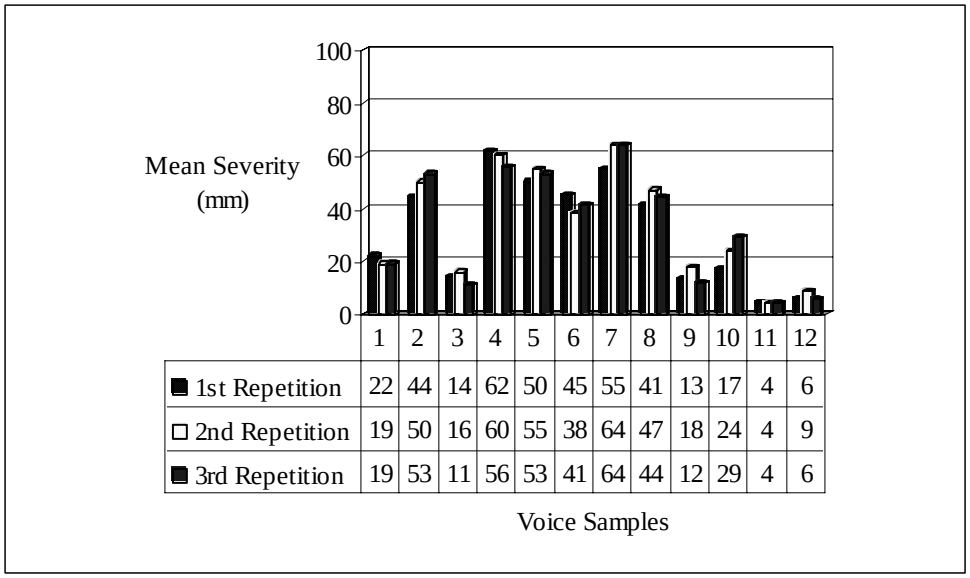
One Standard Deviation in Millimeters for Group 2 Rating of Overall Severity



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	13	15	8	18	19	23	20	16	11	9	6	9
2nd Repetition	12	11	10	12	12	11	17	12	11	16	6	7
3rd Repetition	16	11	9	16	16	16	15	15	10	14	4	7

Table 5

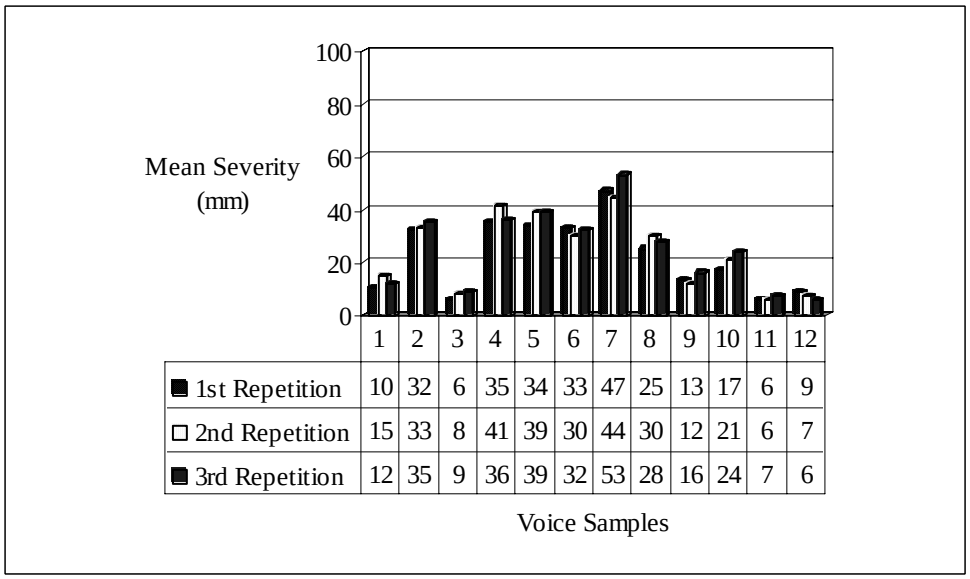
Mean Severity in Millimeters for Group 1 Rating of Roughness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	22	44	14	62	50	45	55	41	13	17	4	6
2nd Repetition	19	50	16	60	55	38	64	47	18	24	4	9
3rd Repetition	19	53	11	56	53	41	64	44	12	29	4	6

Table 6

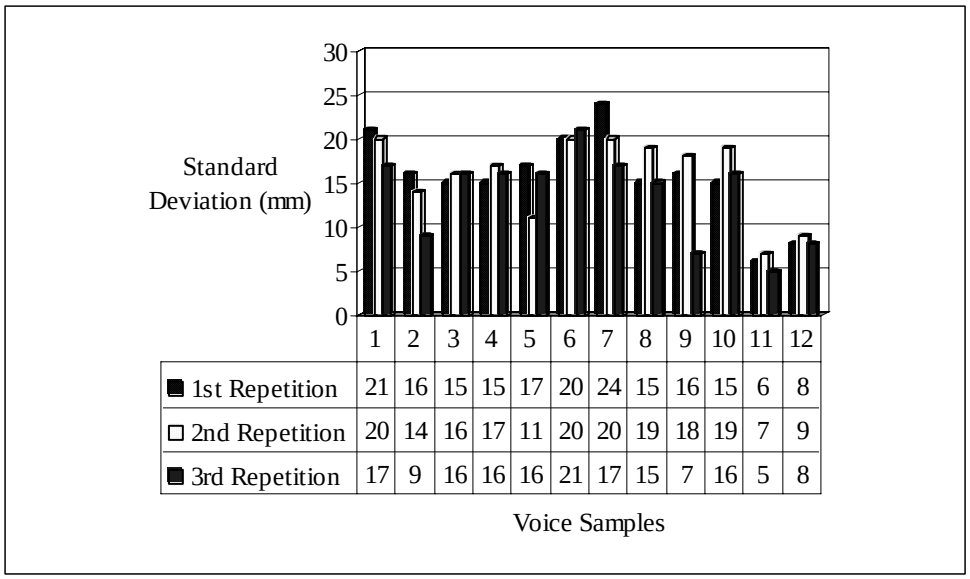
Mean Severity in Millimeters for Group 2 Rating of Roughness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	10	32	6	35	34	33	47	25	13	17	6	9
2nd Repetition	15	33	8	41	39	30	44	30	12	21	6	7
3rd Repetition	12	35	9	36	39	32	53	28	16	24	7	6

Table 7

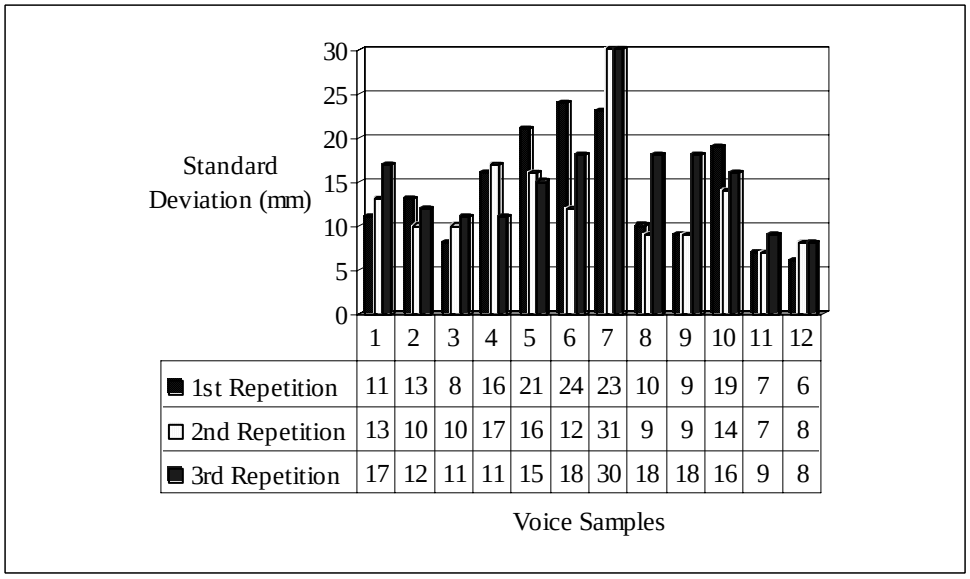
One Standard Deviation in Millimeters for Group 1 Rating of Roughness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	21	16	15	15	17	20	24	15	16	15	6	8
2nd Repetition	20	14	16	17	11	20	20	19	18	19	7	9
3rd Repetition	17	9	16	16	16	21	17	15	7	16	5	8

Table 8

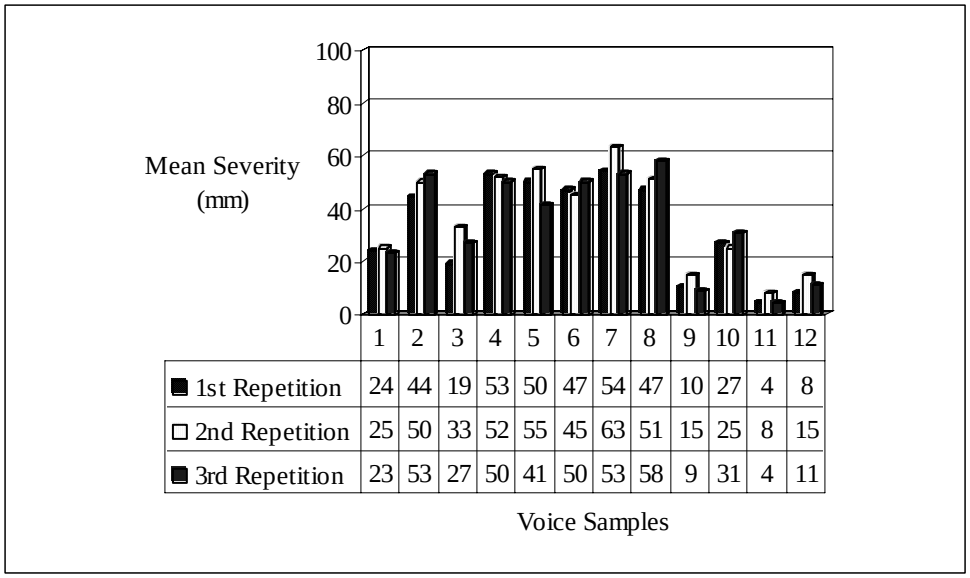
One Standard Deviation in Millimeters for Group 2 Rating of Roughness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	11	13	8	16	21	24	23	10	9	19	7	6
2nd Repetition	13	10	10	17	16	12	31	9	9	14	7	8
3rd Repetition	17	12	11	11	15	18	30	18	18	16	9	8

Table 9

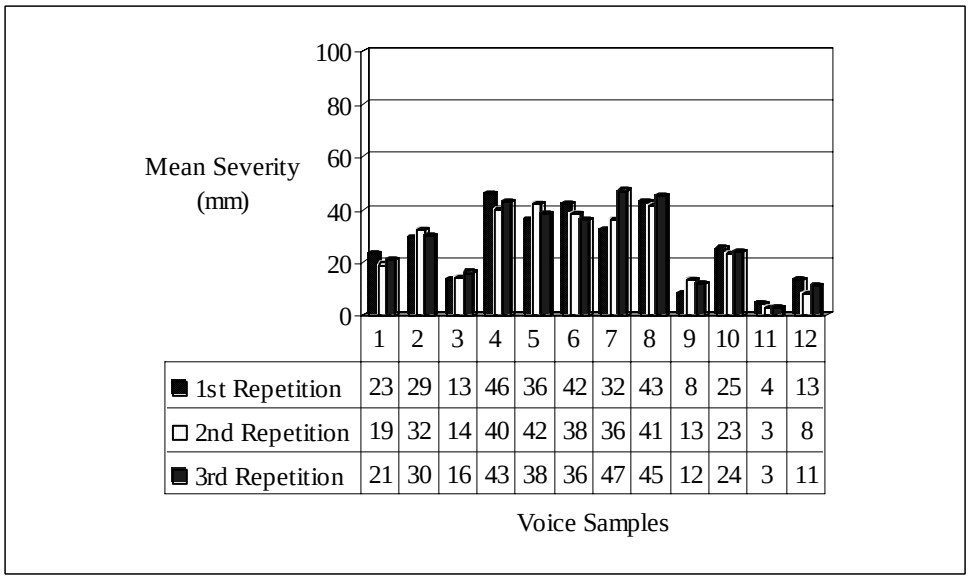
Mean Severity in Millimeters for Group 1 Rating of Breathiness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	24	44	19	53	50	47	54	47	10	27	4	8
2nd Repetition	25	50	33	52	55	45	63	51	15	25	8	15
3rd Repetition	23	53	27	50	41	50	53	58	9	31	4	11

Table 10

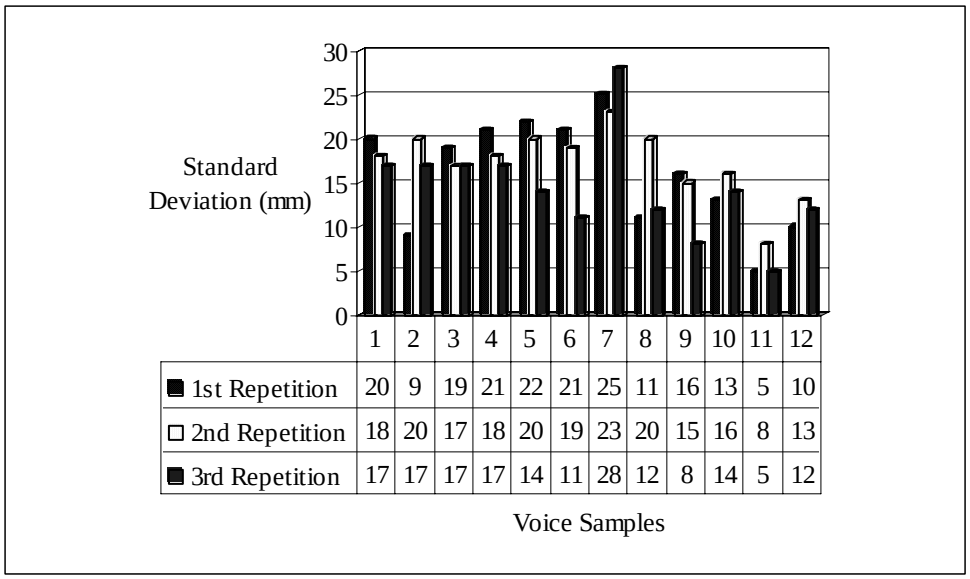
Mean Severity in Millimeters for Group 2 Rating of Breathiness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	23	29	13	46	36	42	32	43	8	25	4	13
2nd Repetition	19	32	14	40	42	38	36	41	13	23	3	8
3rd Repetition	21	30	16	43	38	36	47	45	12	24	3	11

Table 11

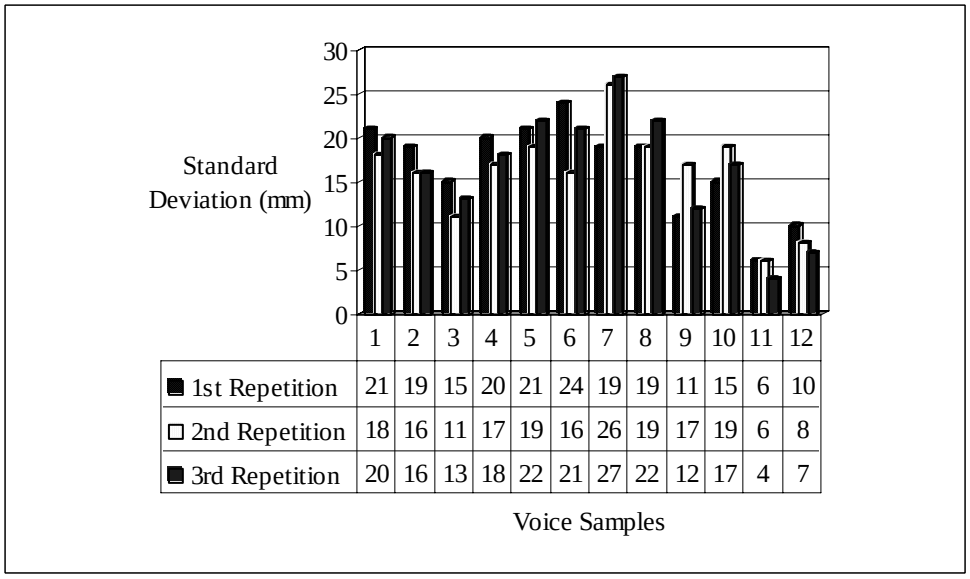
One Standard Deviation in Millimeters for Group 1 Rating of Breathiness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	20	9	19	21	22	21	25	11	16	13	5	10
2nd Repetition	18	20	17	18	20	19	23	20	15	16	8	13
3rd Repetition	17	17	17	17	14	11	28	12	8	14	5	12

Table 12

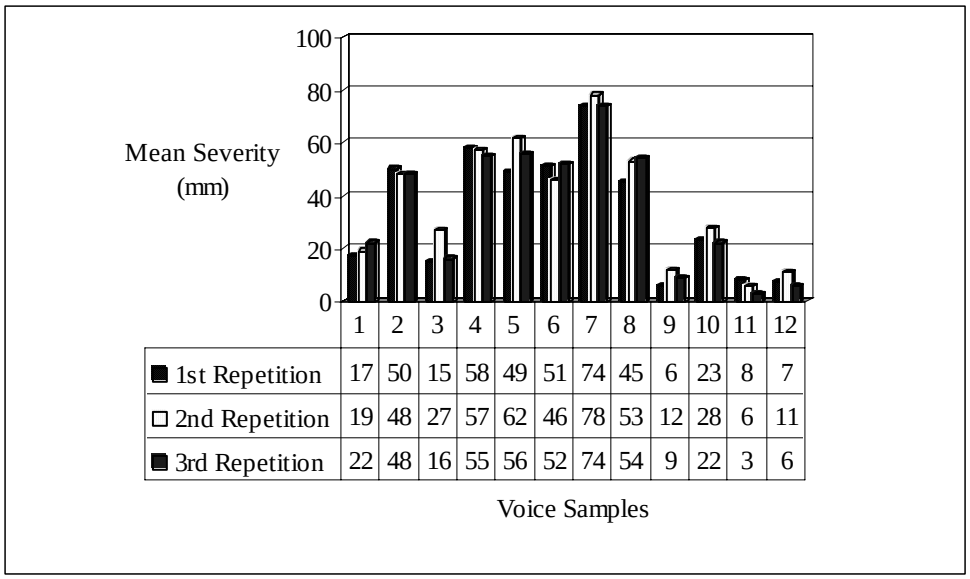
One Standard Deviation in Millimeters for Group 2 Rating of Breathiness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	21	19	15	20	21	24	19	19	11	15	6	10
2nd Repetition	18	16	11	17	19	16	26	19	17	19	6	8
3rd Repetition	20	16	13	18	22	21	27	22	12	17	4	7

Table 13

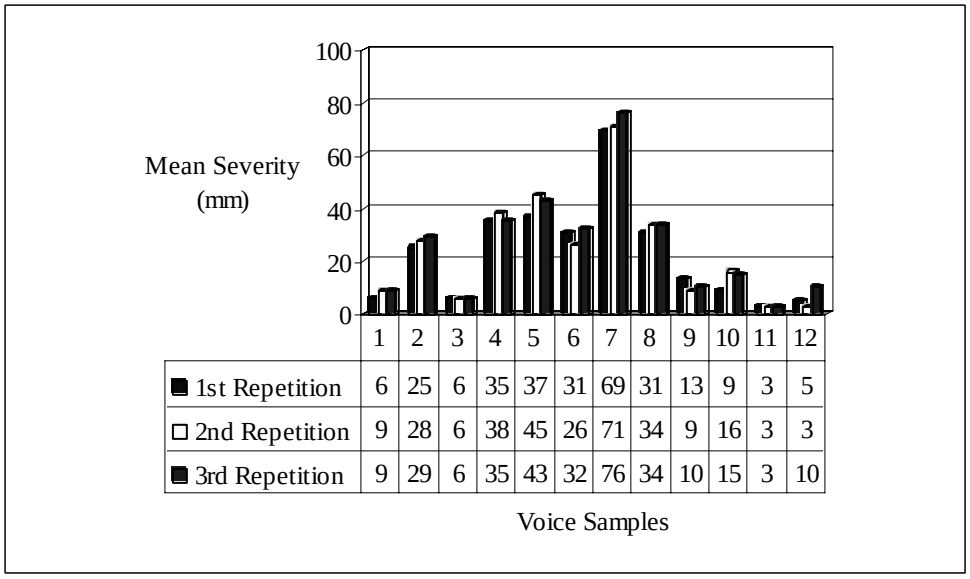
Mean Severity in Millimeters for Group 1 Rating of Strain



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	17	50	15	58	49	51	74	45	6	23	8	7
2nd Repetition	19	48	27	57	62	46	78	53	12	28	6	11
3rd Repetition	22	48	16	55	56	52	74	54	9	22	3	6

Table 14

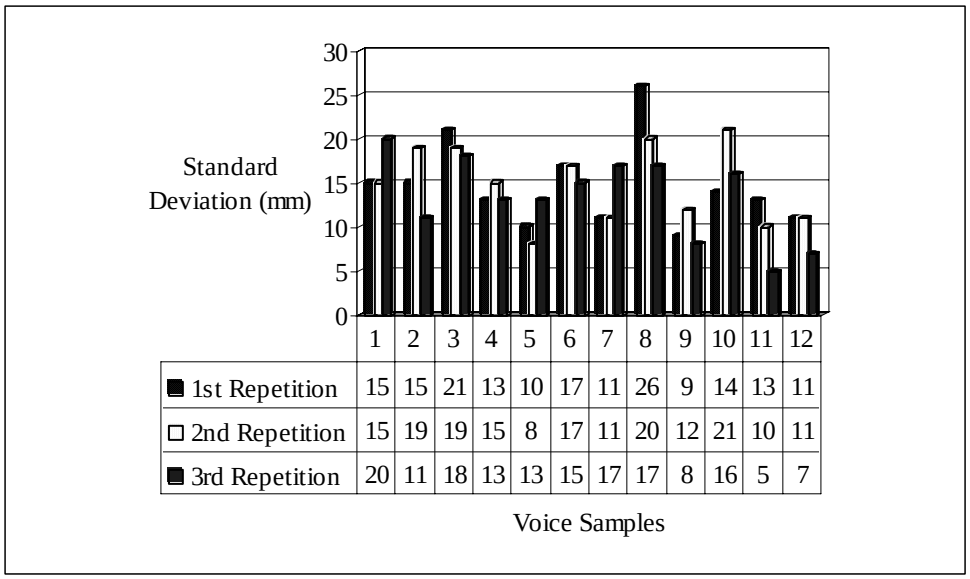
Mean Severity in Millimeters for Group 2 Rating of Strain



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	6	25	6	35	37	31	69	31	13	9	3	5
2nd Repetition	9	28	6	38	45	26	71	34	9	16	3	3
3rd Repetition	9	29	6	35	43	32	76	34	10	15	3	10

Table 15

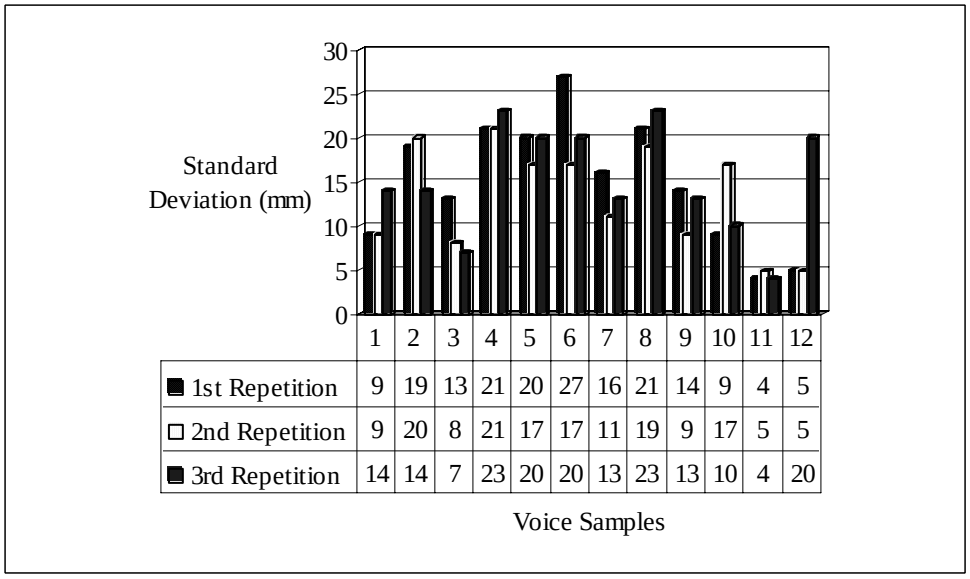
One Standard Deviation in Millimeters for Group 1 Rating of Strain



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	15	15	21	13	10	17	11	26	9	14	13	11
2nd Repetition	15	19	19	15	8	17	11	20	12	21	10	11
3rd Repetition	20	11	18	13	13	15	17	17	8	16	5	7

Table 16

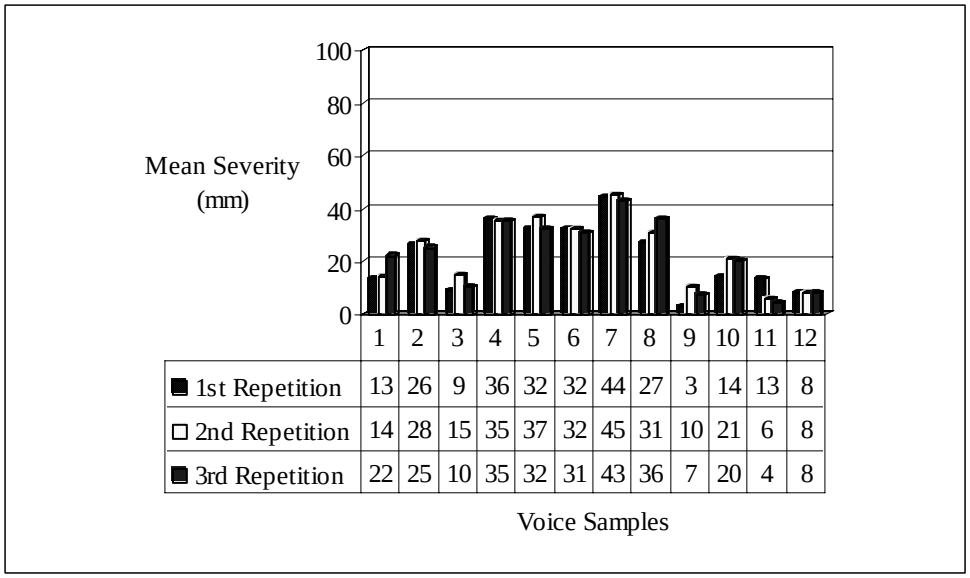
One Standard Deviation in Millimeters for Group 2 Rating of Strain



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	9	19	13	21	20	27	16	21	14	9	4	5
2nd Repetition	9	20	8	21	17	17	11	19	9	17	5	5
3rd Repetition	14	14	7	23	20	20	13	23	13	10	4	20

Table 17

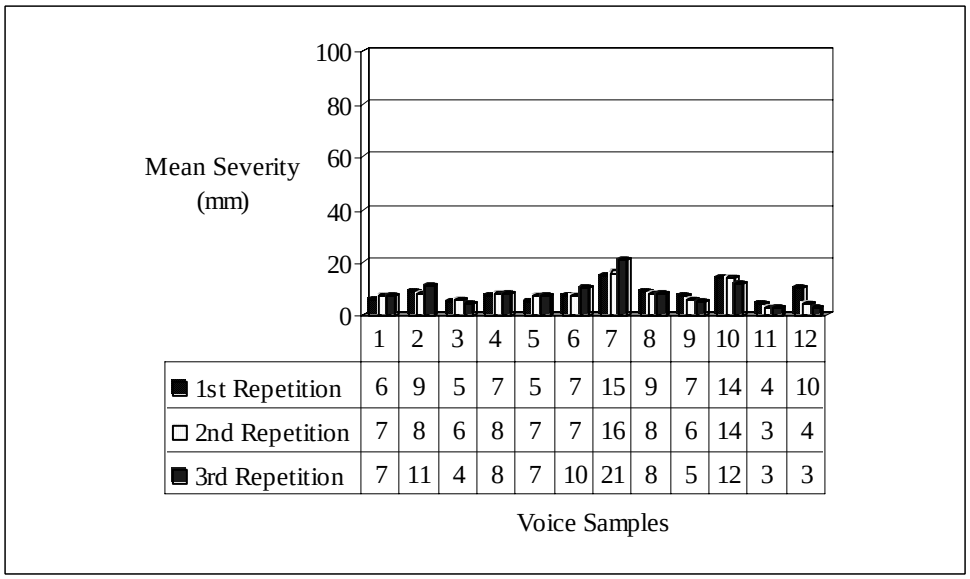
Mean Severity in Millimeters for Group 1 Rating of Pitch



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	13	26	9	36	32	32	44	27	3	14	13	8
2nd Repetition	14	28	15	35	37	32	45	31	10	21	6	8
3rd Repetition	22	25	10	35	32	31	43	36	7	20	4	8

Table 18

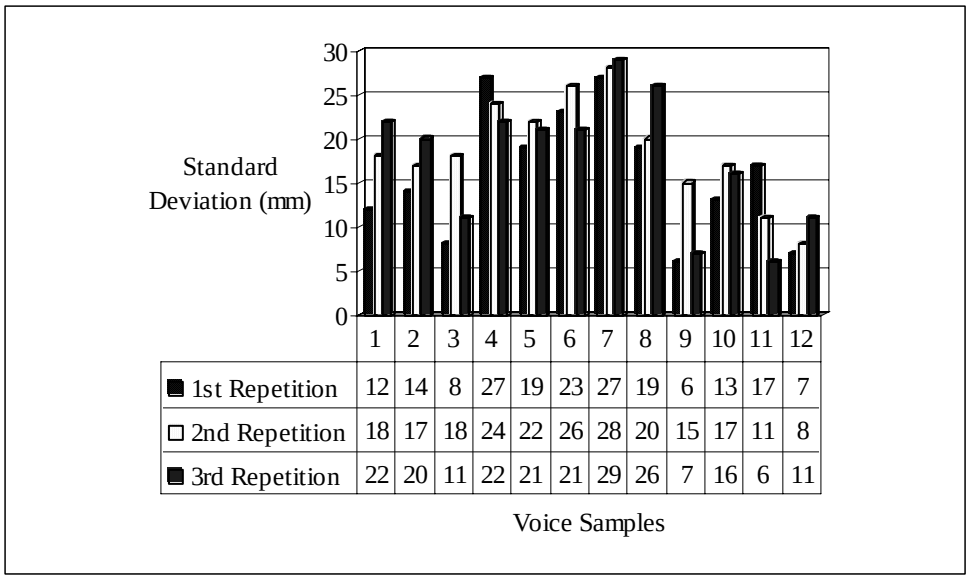
Mean Severity in Millimeters for Group 2 Rating of Pitch



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	6	9	5	7	5	7	15	9	7	14	4	10
2nd Repetition	7	8	6	8	7	7	16	8	6	14	3	4
3rd Repetition	7	11	4	8	7	10	21	8	5	12	3	3

Table 19

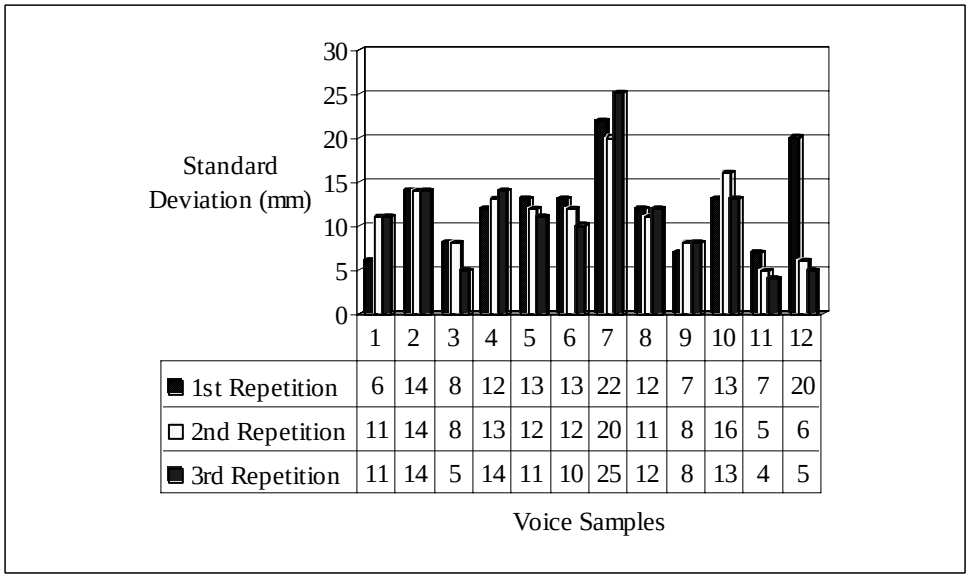
One Standard Deviation in Millimeters for Group 1 Rating of Pitch



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	12	14	8	27	19	23	27	19	6	13	17	7
2nd Repetition	18	17	18	24	22	26	28	20	15	17	11	8
3rd Repetition	22	20	11	22	21	21	29	26	7	16	6	11

Table 20

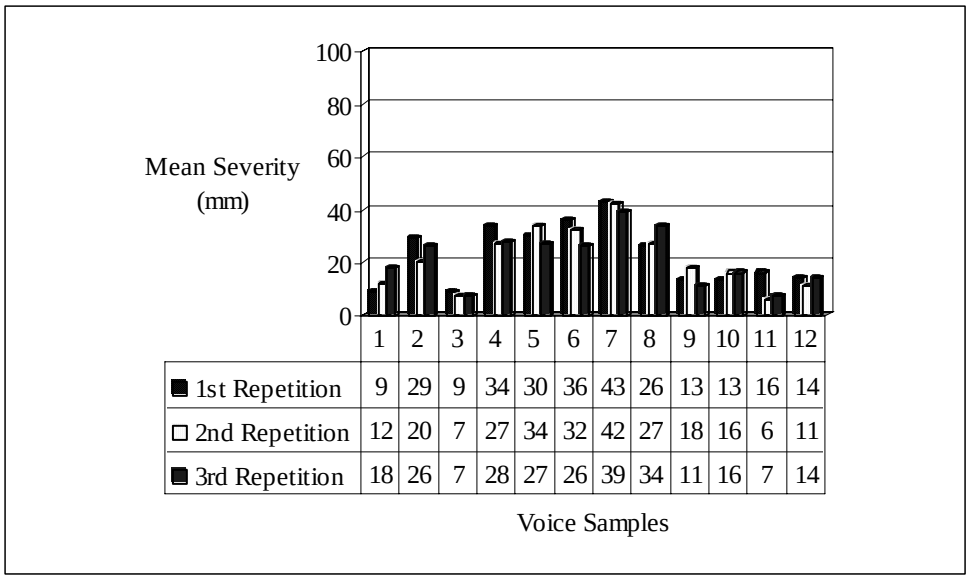
One Standard Deviation in Millimeters for Group 2 Rating of Pitch



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	6	14	8	12	13	13	22	12	7	13	7	20
2nd Repetition	11	14	8	13	12	12	20	11	8	16	5	6
3rd Repetition	11	14	5	14	11	10	25	12	8	13	4	5

Table 21

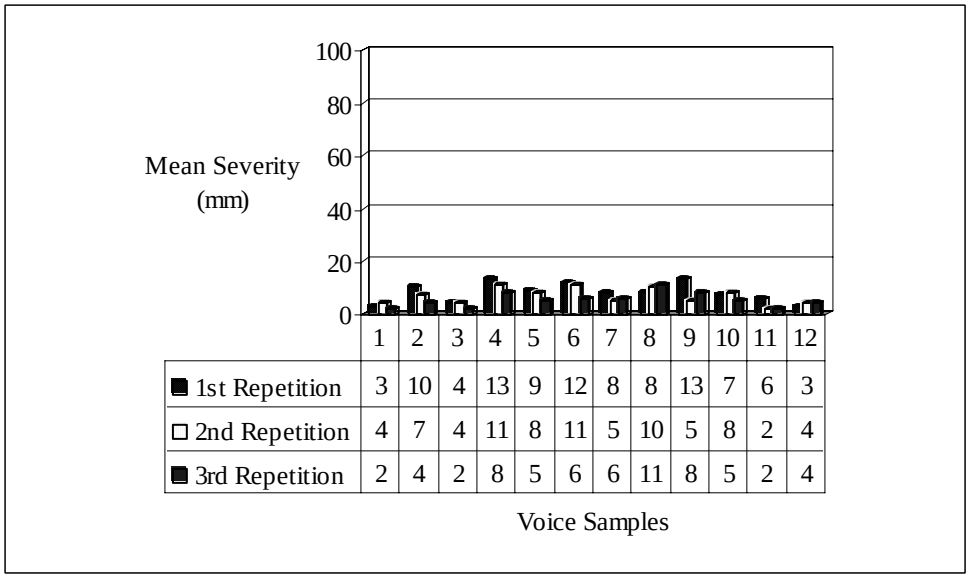
Mean Severity in Millimeters for Group 1 Rating of Loudness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	9	29	9	34	30	36	43	26	13	13	16	14
2nd Repetition	12	20	7	27	34	32	42	27	18	16	6	11
3rd Repetition	18	26	7	28	27	26	39	34	11	16	7	14

Table 22

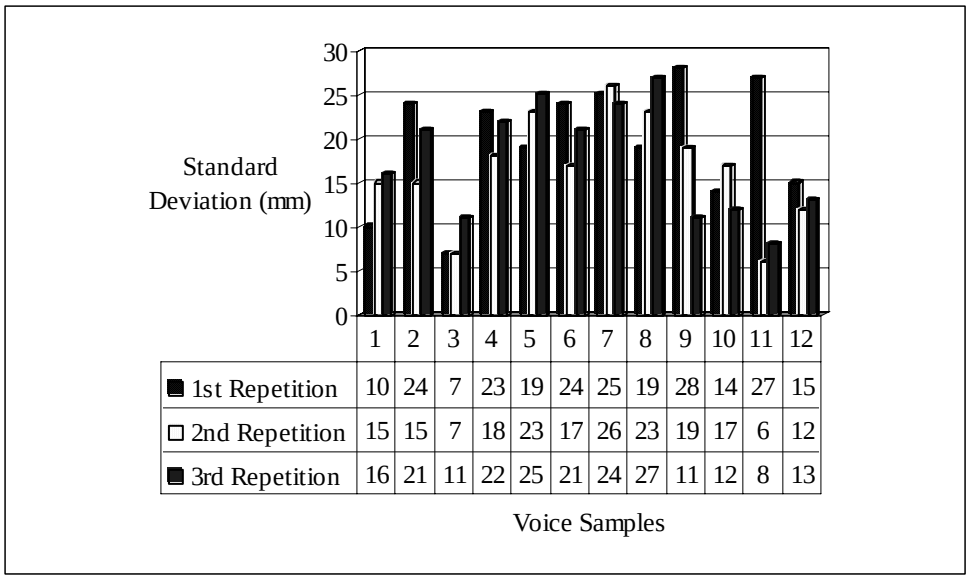
Mean Severity in Millimeters for Group 2 Rating of Loudness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	3	10	4	13	9	12	8	8	13	7	6	3
2nd Repetition	4	7	4	11	8	11	5	10	5	8	2	4
3rd Repetition	2	4	2	8	5	6	6	11	8	5	2	4

Table 23

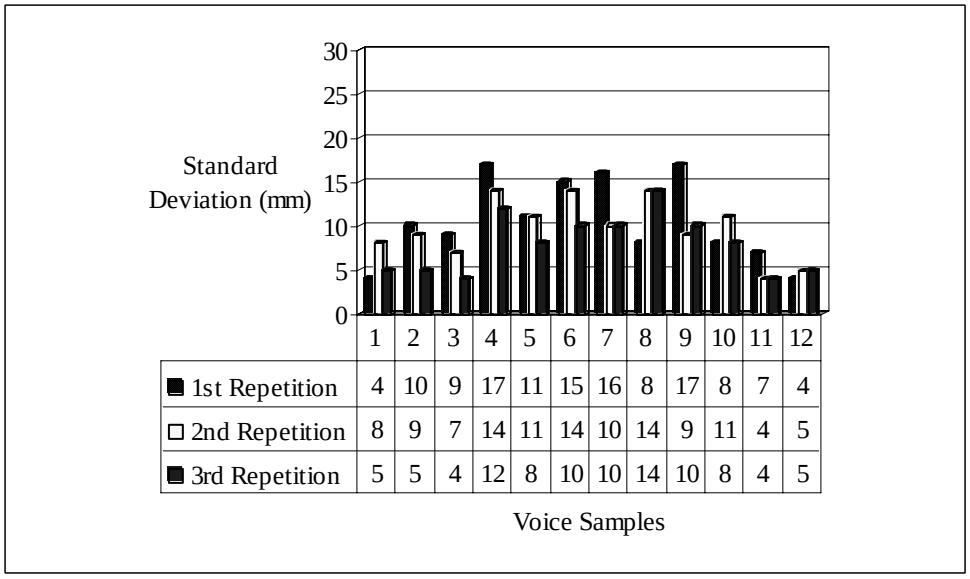
One Standard Deviation in Millimeters for Group 1 Rating of Loudness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	10	24	7	23	19	24	25	19	28	14	27	15
2nd Repetition	15	15	7	18	23	17	26	23	19	17	6	12
3rd Repetition	16	21	11	22	25	21	24	27	11	12	8	13

Table 24

One Standard Deviation in Millimeters for Group 2 Rating of Loudness



	1	2	3	4	5	6	7	8	9	10	11	12
1st Repetition	4	10	9	17	11	15	16	8	17	8	7	4
2nd Repetition	8	9	7	14	11	14	10	14	9	11	4	5
3rd Repetition	5	5	4	12	8	10	10	14	10	8	4	5

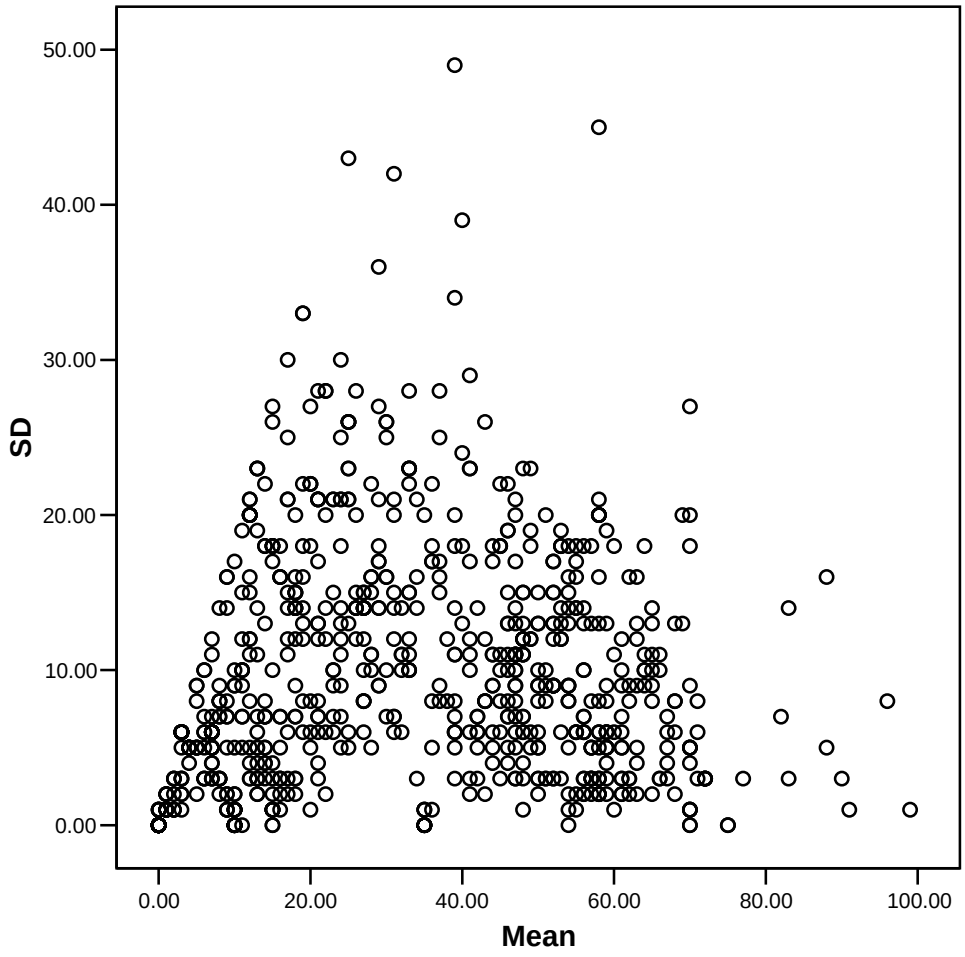


Figure 1: Mean and Standard Deviation for Each Participant's Rating Over the 3 Repetitions for Group 1; Including all Perceptual Indices

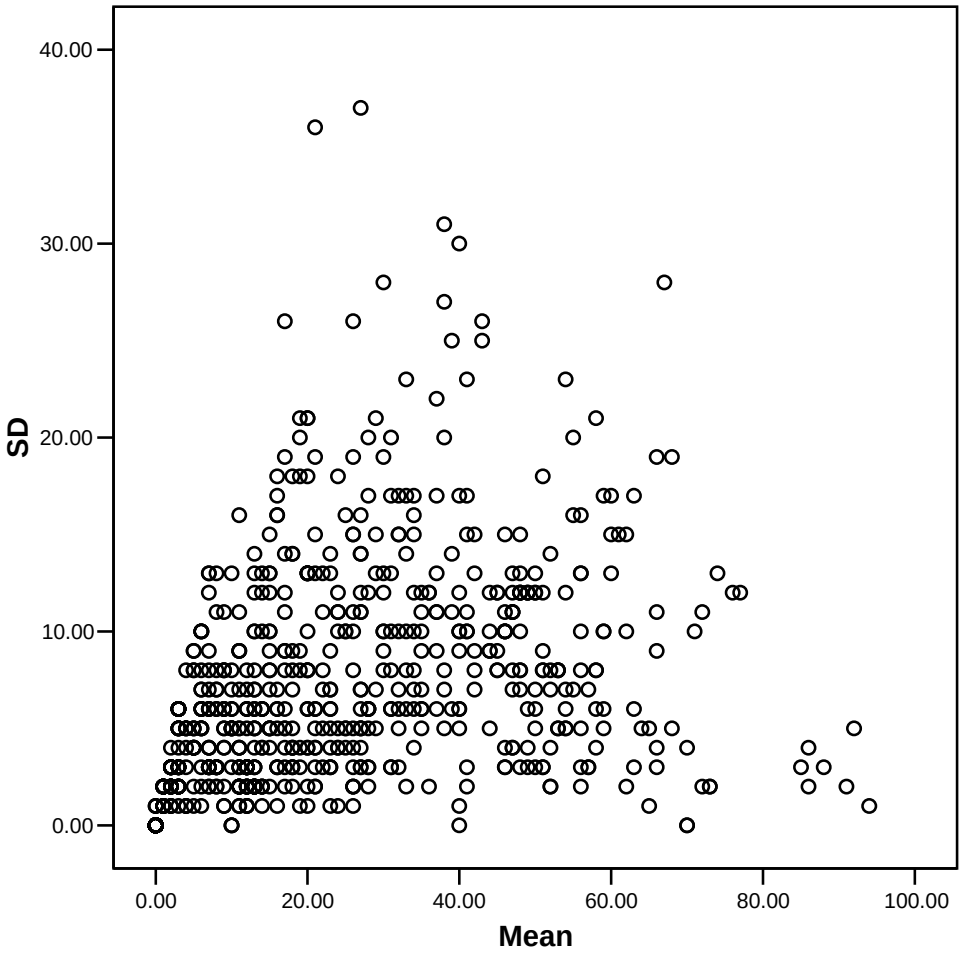


Figure 2: Mean and Standard Deviation for Each Participant's Rating Over the 3 Repetitions for Group 2; Including all Perceptual Indices

Table 25

Comparison of Group 1 and Group 2 Mean Scores for Overall Severity

Voice Samples	*p*	
1	F = .290,	*p* = .595
2	**F = 12.290,**	***p* = .002**
3	F = 1.566,	*p* = .224
4	**F = 10.34,**	***p* = .004**
5	F = 2.711,	*p* = .114
6	**F = 5.170,**	***p* = .033**
7	F = .110,	*p* = .744
8	**F = 5.745,**	***p* = .025**
9	F = .167,	*p* = .687
10	F = .645,	*p* = .430
11	F = .790,	*p* = .384
12	F = .019,	*p* = .893

***p* < .05: significance**

Table 26

Comparison of Group 1 and Group 2 Mean Scores for Roughness

Voice Samples	*p*	
1	**F = 9.721,**	***p* = .005**
2	F = .027,	*p* = .872
3	F = 1.943,	*p* = .177
4	**F = 10.711,**	***p* = .003**
5	F = 1.163,	*p* = .293
6	**F = 9.004,**	***p* = .007**
7	F = .052,	*p* = .052
8	F = .033,	*p* = .857
9	F = 1.564,	*p* = .224
10	F = .944,	*p* = .342
11	F = .285,	*p* = .598
12	F = 4.040,	*p* = .057

***p* < .05: significance**

Table 27

Comparison of Group 1 and Group 2 Mean Scores for Breathiness

Voice Samples	*p*	
1	F = .195,	*p* = .663
2	F = 2.761,	*p* = .111
3	**F = 5.206,**	***p* = .033**
4	F = 1.734,	*p* = .202
5	F = 1.828,	*p* = .190
6	F = 1.601,	*p* = .219
7	F = 4.061,	*p* = .056
8	F = 2.080,	*p* = .163
9	F = .006,	*p* = .941
10	F = .426,	*p* = .521
11	F = .000,	*p* = 1.000
12	F = .001,	*p* = .981

***p* < .05: significance**

Table 28

Comparison of Group 1 and Group 2 Mean Scores for Strain

Voice Samples	*p*	
1	**F = 10.509,**	***p* = .004**
2	**F = 17.642,**	***p* = .000**
3	**F = 15.488,**	***p* = .001**
4	**F = 10.801,**	***p* = .003**
5	**F = 8.193,**	***p* = .009**
6	**F = 9.299,**	***p* = .006**
7	F = .345,	*p* = .563
8	**F = 8.430,**	***p* = .008**
9	F = .218,	*p* = .645
10	**F = 4.675,**	***p* = .042**
11	F = 1.208,	*p* = .284
12	F = .330,	*p* = .571

***p* < .05: significance**

Table 29

Comparison of Group 1 and Group 2 Mean Scores for Pitch

Voice Samples	*p*	
1	F = 3.237,	*p* = .086
2	**F = 7.894,**	***p* = .010**
3	F = 2.942,	*p* = .100
4	**F = 12.642,**	***p* = .002**
5	**F = 17.079,**	***p* = .000**
6	**F = 15.913,**	***p* = .001**
7	**F = 7.677,**	***p* = .011**
8	**F = 12.041,**	***p* = .002**
9	F = .058,	*p* = .812
10	F = .924,	*p* = .347
11	F = 3.314,	*p* = .082
12	**F= 4.881,**	***p* = .038**

***p* < .05: significance**

Table 30

Comparison of Group 1 and Group 2 Mean Scores for Loudness

Voice Samples	*p*	
1	**F = 6.621,**	***p* = .017**
2	**F = 16.038,**	***p* = .001**
3	F = 1.663,	*p* = .211
4	**F = 8.360,**	***p* = .008**
5	**F = 13.419,**	***p* = .001**
6	**F = 11.647,**	***p* = .002**
7	**F = 21.266,**	***p* = .000**
8	**F = 7.655,**	***p* = .011**
9	F = 1.035,	*p* = .320
10	F = 4.105,	*p* = .055
11	F = 3.274,	*p* = .084
12	**F = 9.493,**	***p* = .005**

***p* < .05: significance**

Table 31

Correlation Matrix for Significant Pearson's Product Moment Coefficient for Group 1 and Group 2, for all Ratings.

Group	Indice	1st vs. 2nd		1st vs. 3rd		2nd vs. 3rd	
		r	P	r	P	r	P
1	Overall Severity	.747	.000	**.816**	**.000**	.786	.000
2	*Overall Severity*	***.830***	***.000***	*.762*	*.000*	***.907***	***.000***
1	Roughness	.731	.000	.782	.000	**.803**	**.000**
2	*Roughness*	*.789*	*.000*	***.832***	***.000***	***.862***	***.000***
1	Breathiness	.732	.000	.744	.000	.770	.000
2	*Breathiness*	*.786*	*.000*	*.799*	*.000*	***.841***	***.000***
1	Strain	.761	.000	.771	.000	**.810**	**.000**
2	*Strain*	***.869***	***.000***	***.822***	***.000***	***.880***	***.000***
1	Pitch	.687	.000	.767	.000	**.802**	**.000**
2	*Pitch*	***.864***	***.000***	*.717*	*.000*	*.759*	*.000*
1	Loudness	.656	.000	.607	.000	.760	.000
2	*Loudness*	*.686*	*.000*	*.700*	*.000*	*.671*	*.000*

0.0 to 0.2 Very weak to negligible correlation
0.2 to 0.4 Weak, low correlation (not very significant)
0.5 to 0.7 Moderate correlation
0.8 to 0.9 Strong, high correlation
0.9 to 1.0 Very strong correlation

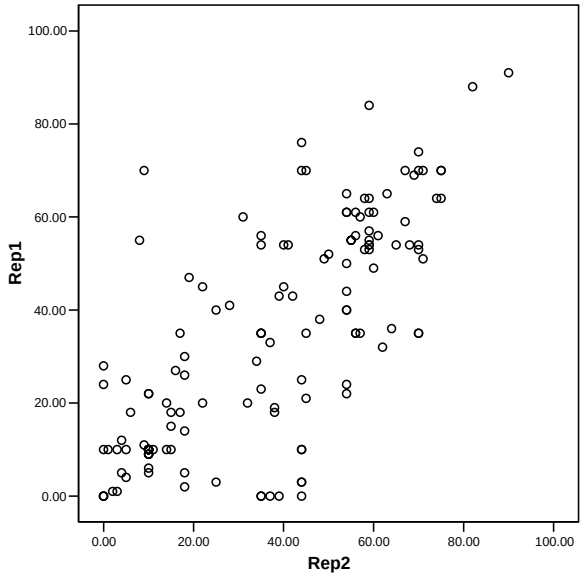


Figure 3: Comparison of the 1st and 2nd Repetitions of Overall Severity in Group 1; Including all Voice Samples.

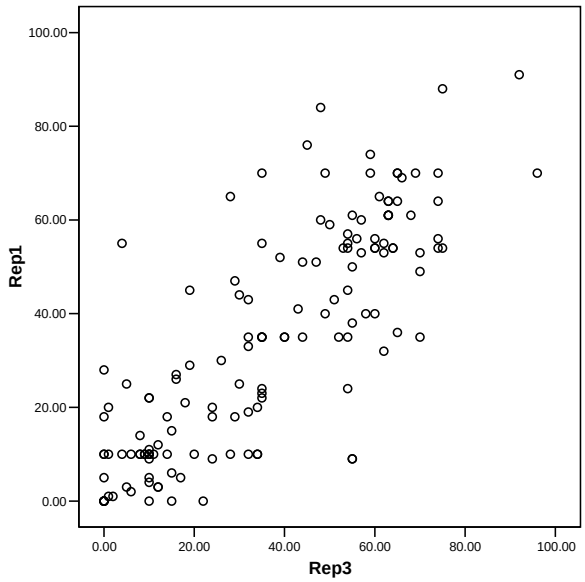


Figure 4: Comparison of the 1st and 3rd Repetitions of Overall Severity in Group 1; Including all Voice Samples.

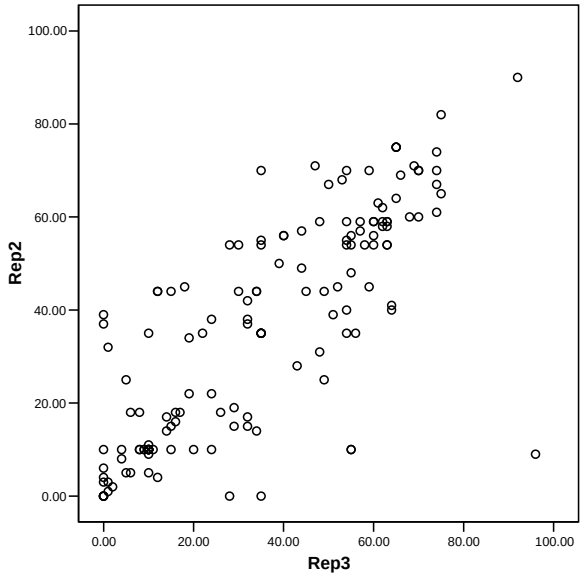


Figure 5: Comparison of the 2nd and 3rd Repetitions of Overall Severity in Group 1; Including all Voice Samples.

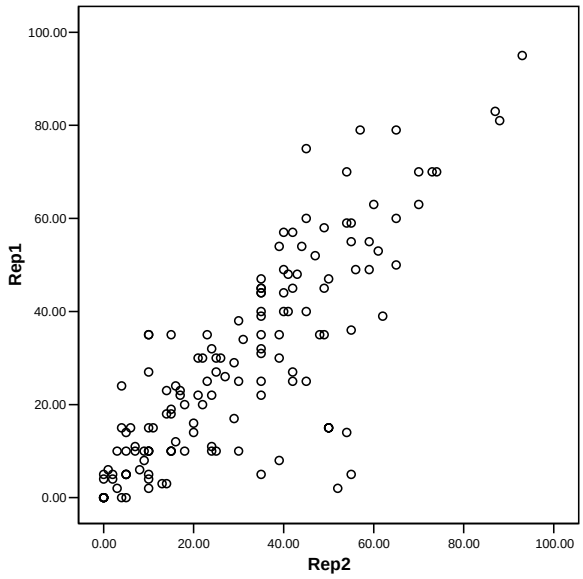


Figure 6: Comparison of the 1st and 2nd Repetitions of Overall Severity in Group 2; Including all Voice Samples.

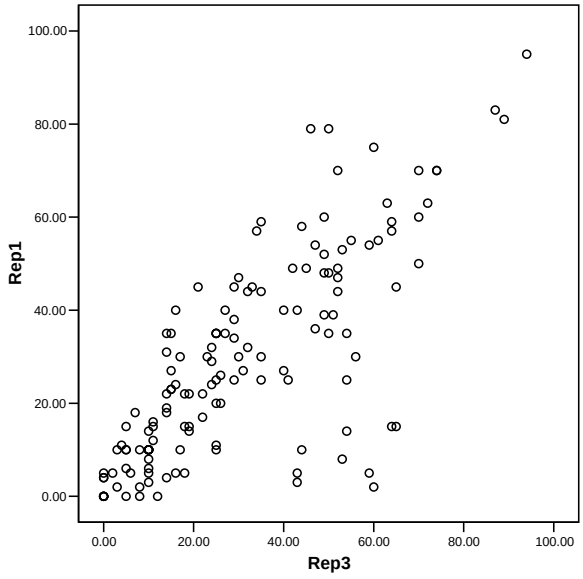


Figure 7: Comparison of the 1st and 3rd Repetitions of Overall Severity in Group 2; Including all Voice Samples.

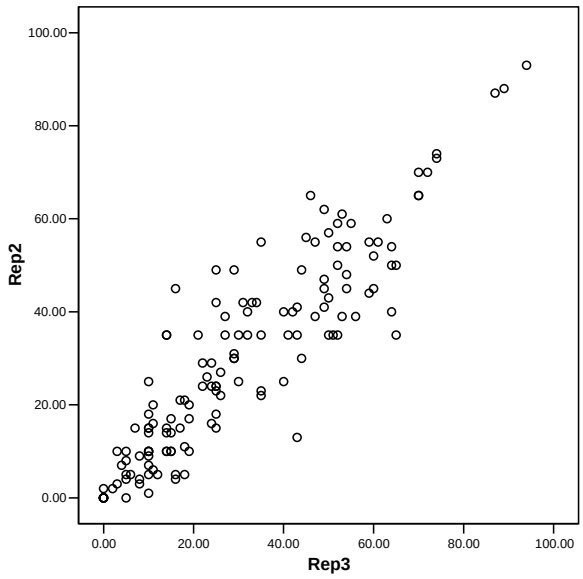


Figure 8: Comparison of the 2nd and 3rd Repetitions of Overall Severity in Group 2; Including all Voice Samples.

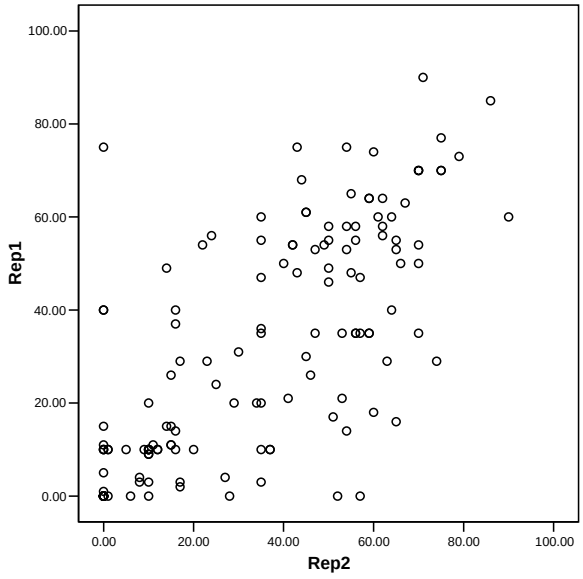


Figure 9: Comparison of the 1st and 2nd Repetitions of Roughness in Group 1; Including all Voice Samples.

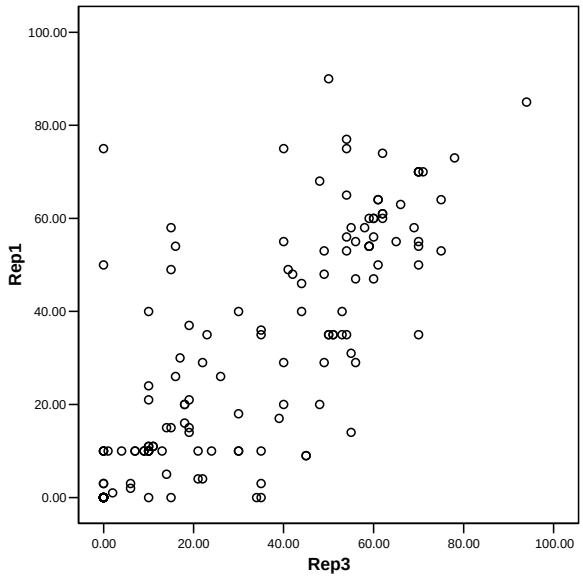


Figure 10: Comparison of the 1st and 3rd Repetitions of Roughness in Group 1; Including all Voice Samples.

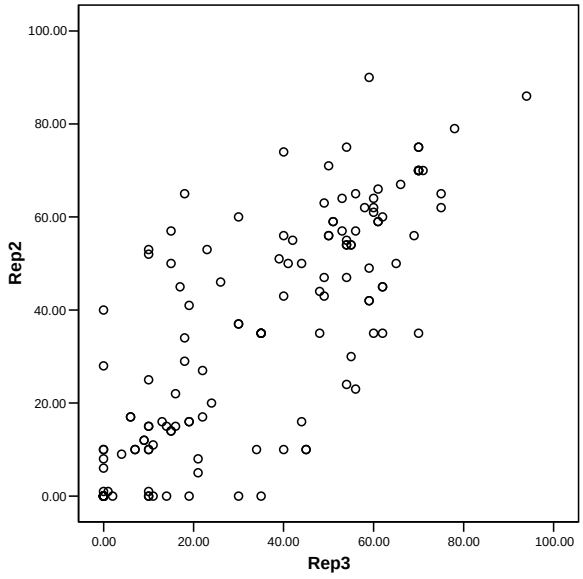


Figure 11: Comparison of the 2nd and 3rd Repetitions of Roughness in Group 1; Including all Voice Samples.

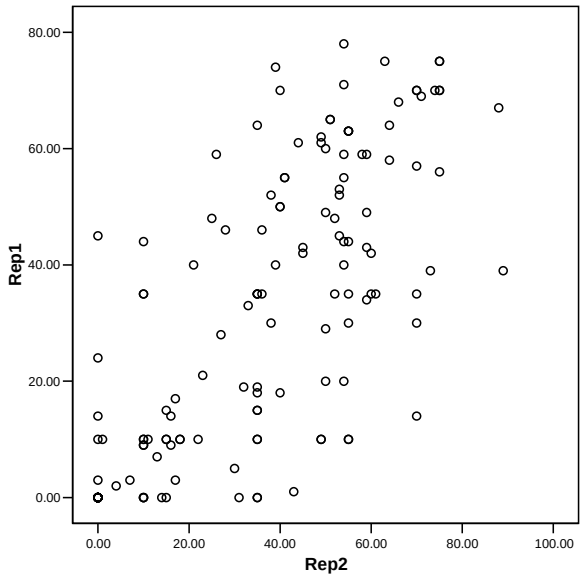


Figure 12: Comparison of the 1st and 2nd Repetitions of Roughness in Group 2; Including all Voice Samples.

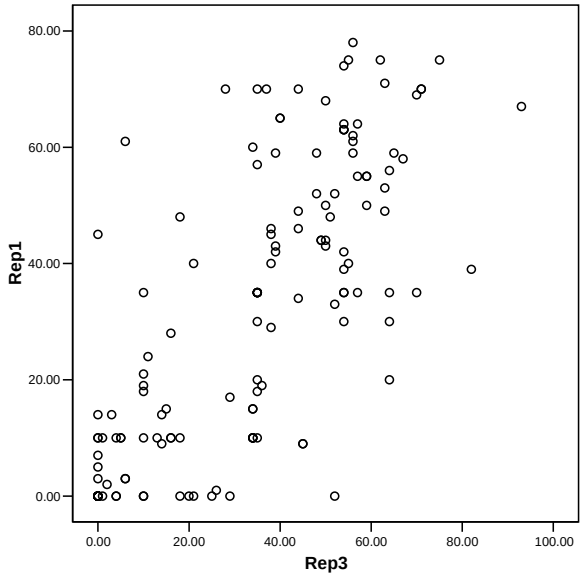


Figure 13: Comparison of the 1st and 3rd Repetitions of Roughness in Group 2; Including all Voice Samples.

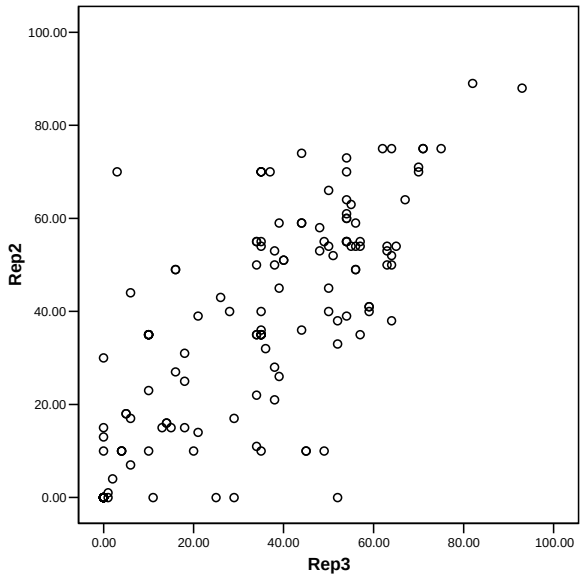


Figure 14: Comparison of the 2nd and 3rd Repetitions of Roughness in Group 2; Including all Voice Samples.

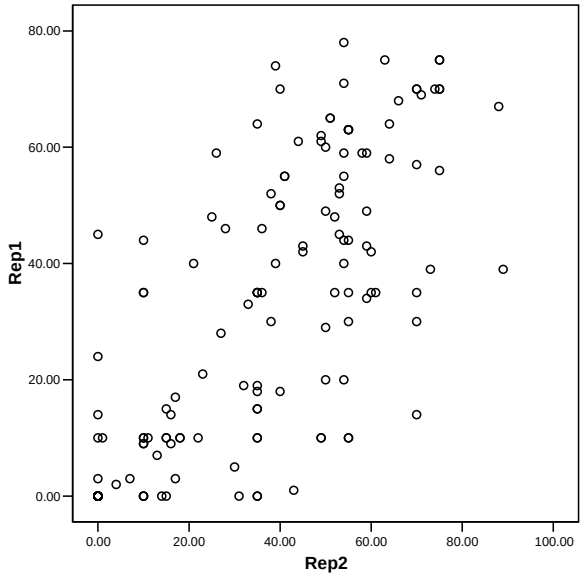


Figure 15: Comparison of the 1st and 2nd Repetitions of Breathiness in Group 1; Including all Voice Samples.

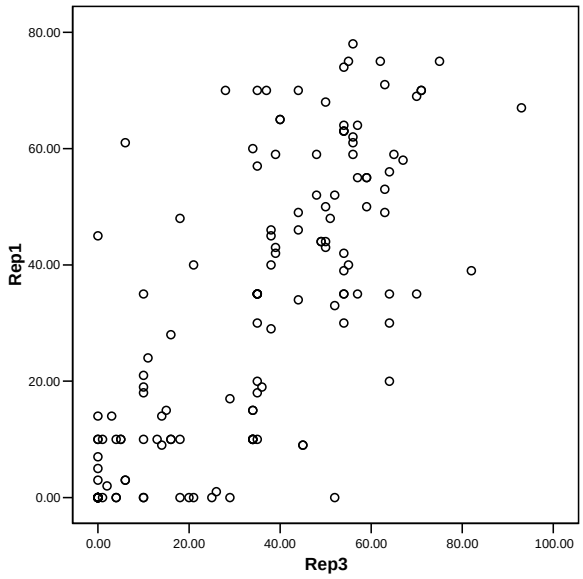


Figure 16: Comparison of the 1st and 3rd Repetitions of Breathiness in Group 1; Including all Voice Samples.

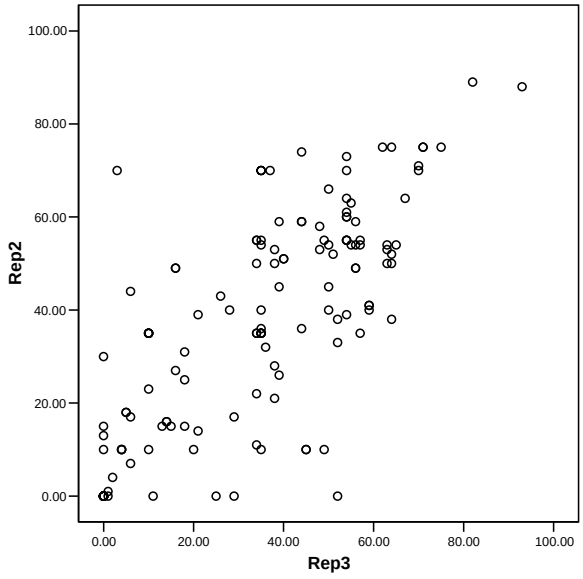


Figure 17: Comparison of the 2nd and 3rd Repetitions of Breathiness in Group 1; Including all Voice Samples.

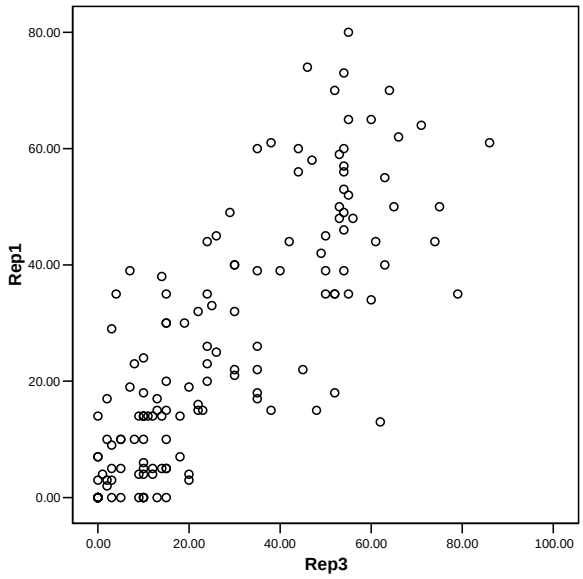


Figure 18: Comparison of the 1st and 2nd Repetitions of Breathiness in Group 2; Including all Voice Samples.

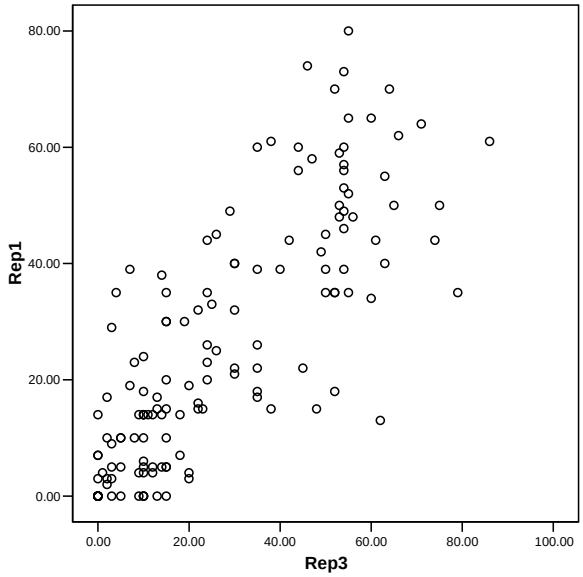


Figure 19: Comparison of the 1st and 3rd Repetitions of Breathiness in Group 2; Including all Voice Samples.

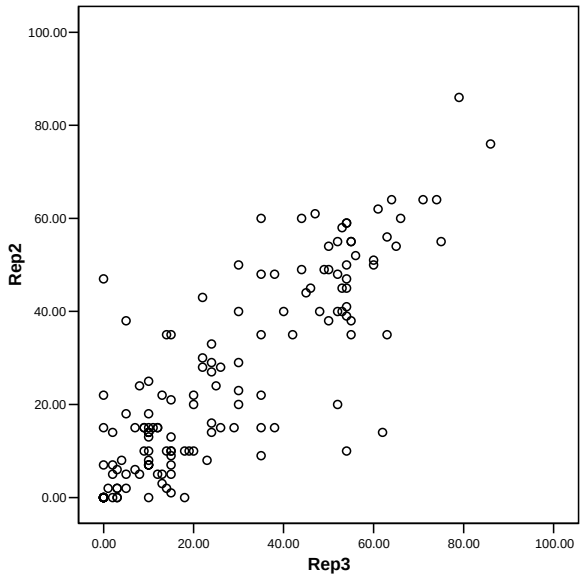


Figure 20: Comparison of the 2nd and 3rd Repetitions of Breathiness in Group 2; Including all Voice Samples.

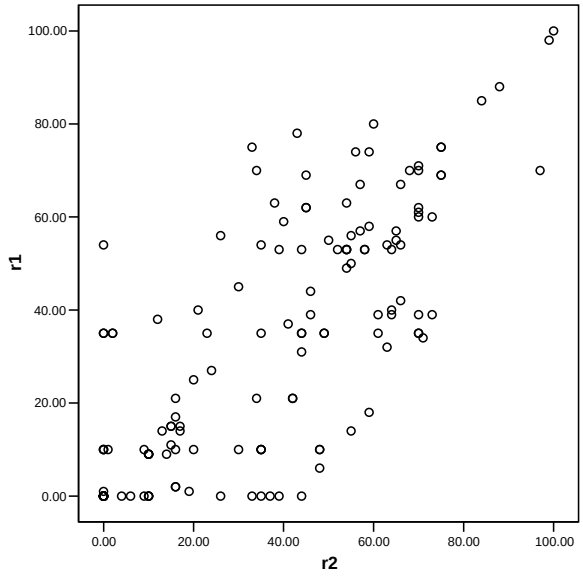


Figure 21: Comparison of the 1st and 2nd Repetitions of Strain in Group 1; Including all Voice Samples.

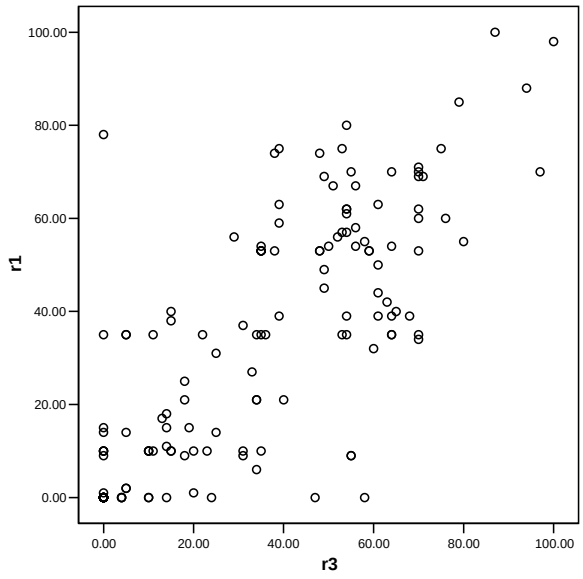


Figure 22: Comparison of the 1st and 3rd Repetitions of Strain in Group 1; Including all Voice Samples.

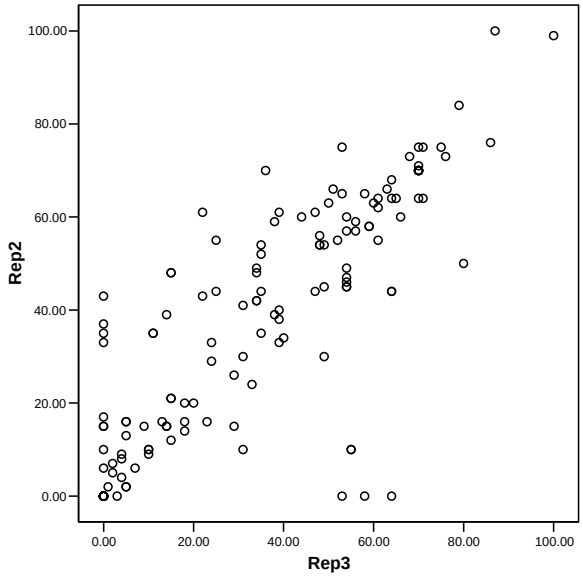


Figure 23: Comparison of the 2nd and 3rd Repetitions of Strain in Group 1; Including all Voice Samples.

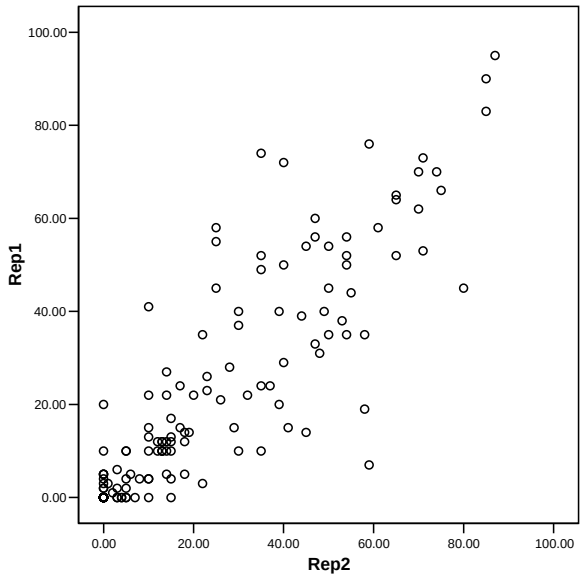


Figure 24: Comparison of the 1st and 2nd Repetitions of Strain in Group 2; Including all Voice Samples.

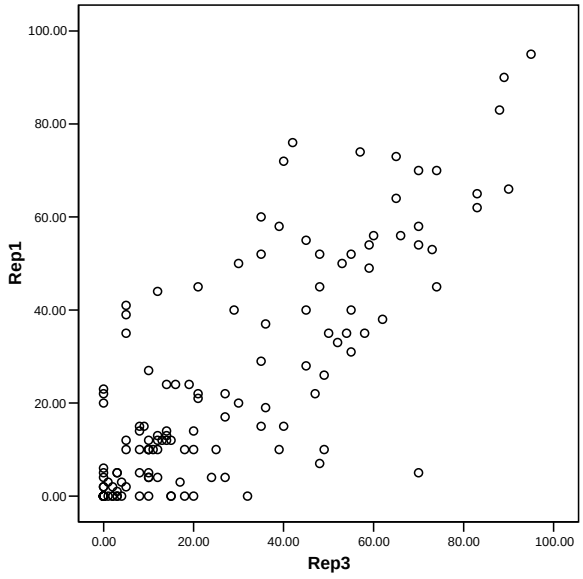


Figure 25: Comparison of the 1st and 3rd Repetitions of Strain in Group 2; Including all Voice Samples.

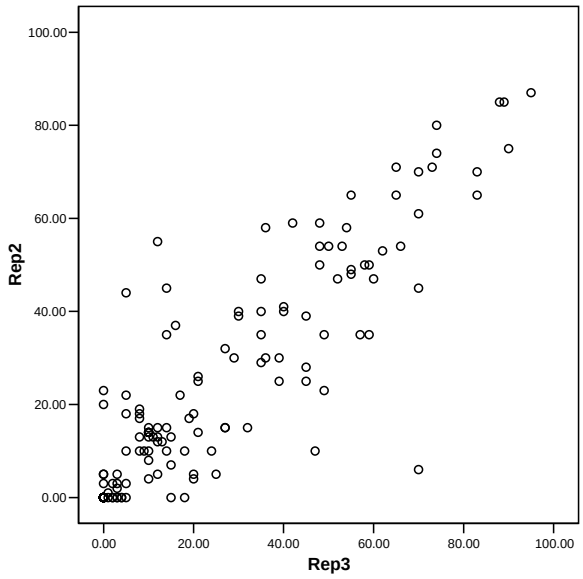


Figure 26: Comparison of the 2nd and 3rd Repetitions of Strain in Group 2; Including all Voice Samples.

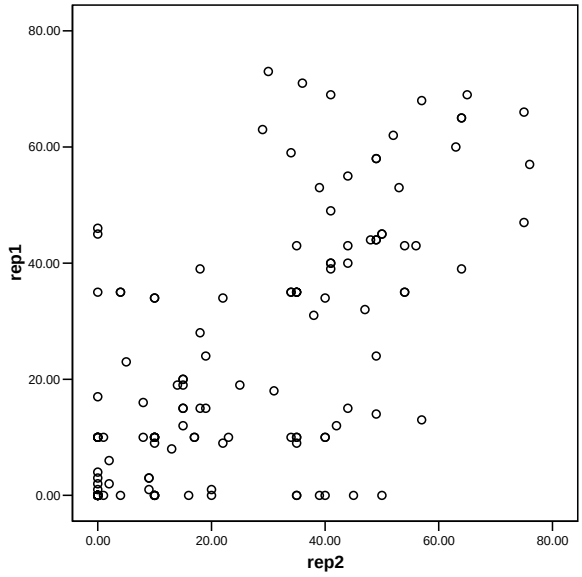


Figure 27: Comparison of the 1st and 2nd Repetitions of Pitch in Group 1; Including all Voice Samples.

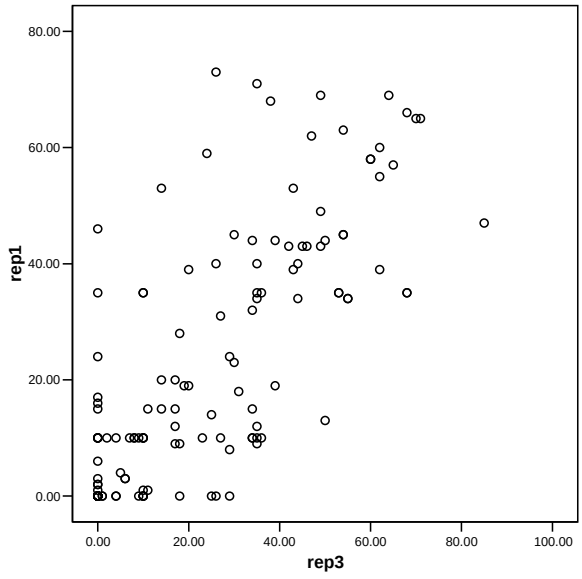


Figure 28: Comparison of the 1st and 3rd Repetitions of Pitch in Group 1; Including all Voice Samples.

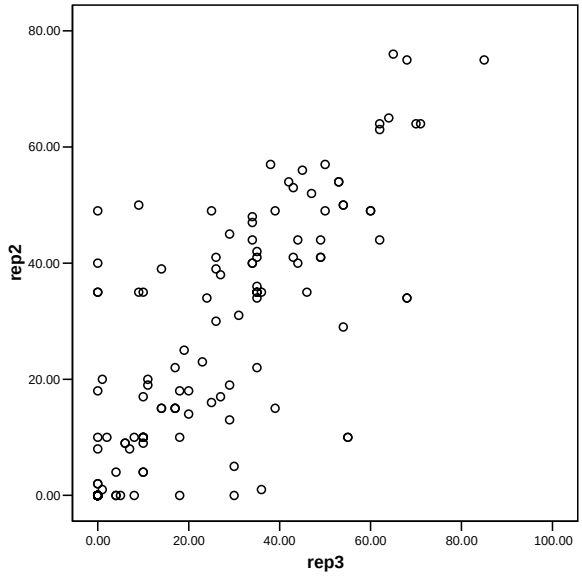


Figure 29: Comparison of the 2nd and 3rd Repetitions of Pitch in Group 1; Including all Voice Samples.

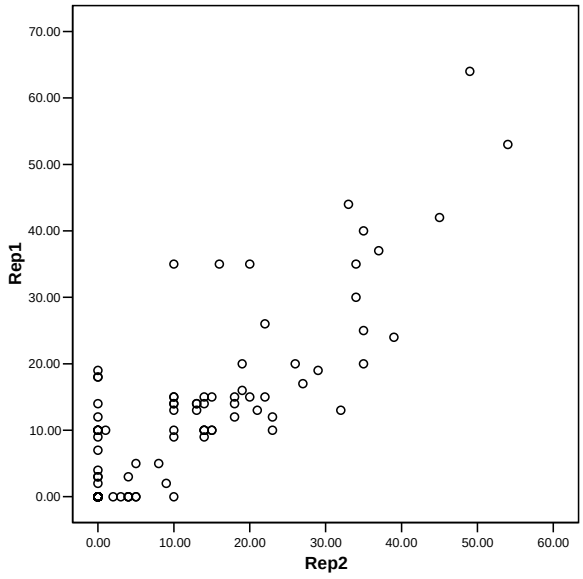


Figure 30: Comparison of the 1st and 2nd Repetitions of Pith in Group 2; Including all Voice Samples.

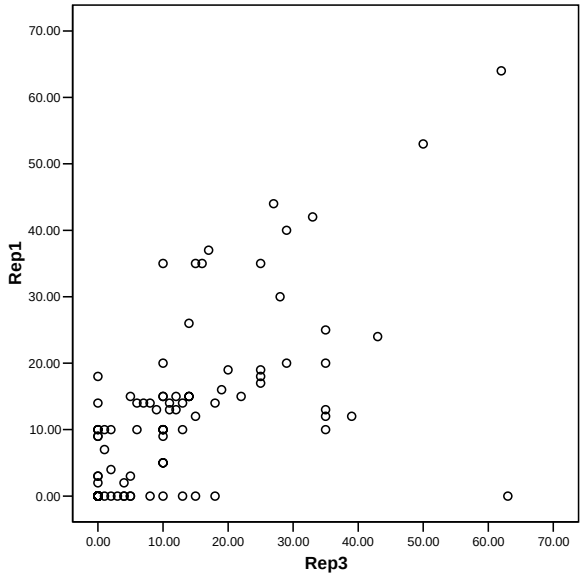


Figure 31: Comparison of the 1st and 3rd Repetitions of Pitch in Group 2; Including all Voice Samples.

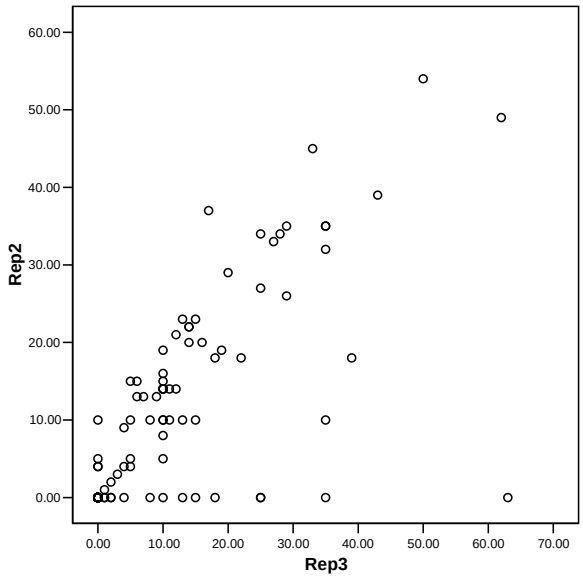


Figure 32: Comparison of the 2nd and 3rd Repetitions of Pitch in Group 2; Including all Voice Samples.

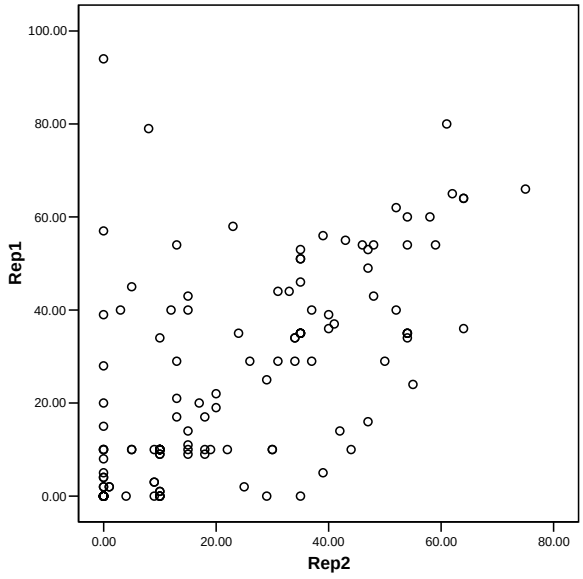


Figure 33: Comparison of the 1st and 2nd Repetitions of Loudness in Group 1; Including all Voice Samples.

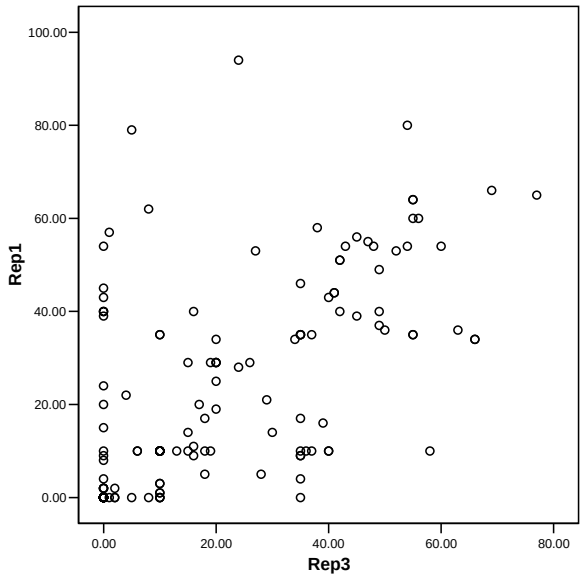


Figure 34: Comparison of the 1st and 3rd Repetitions of Loudness in Group 1; Including all Voice Samples.

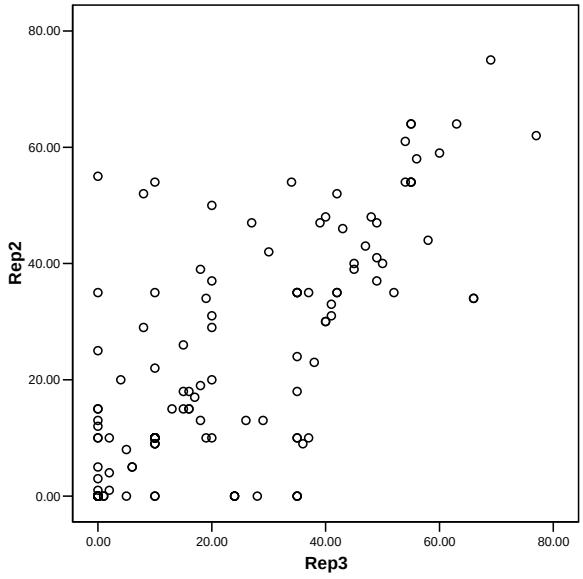


Figure 35: Comparison of the 2nd and 3rd Repetitions of Loudness in Group 1; Including all Voice Samples.

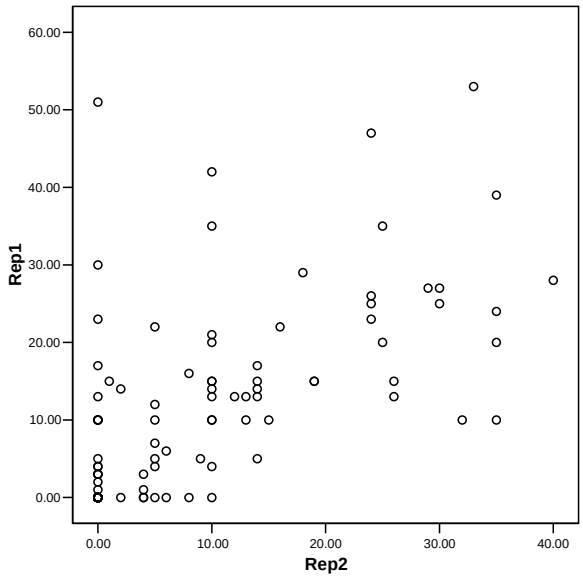


Figure 36: Comparison of the 1st and 2nd Repetitions of Loudness in Group 2; Including all Voice Samples.

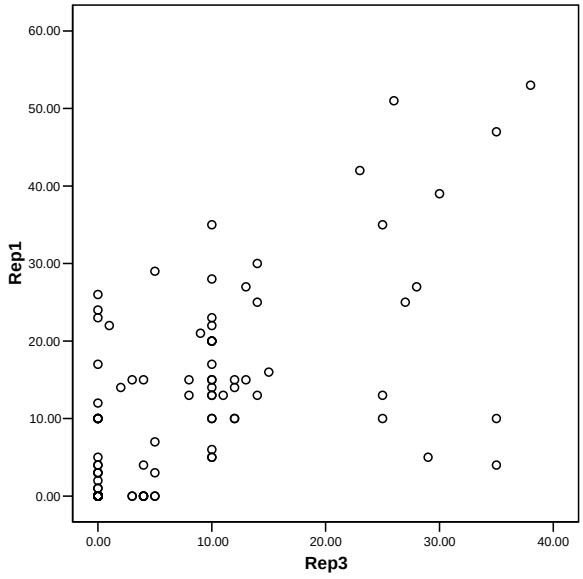


Figure 37: Comparison of the 1st and 3rd Repetitions of Loudness in Group 2; Including all Voice Samples.

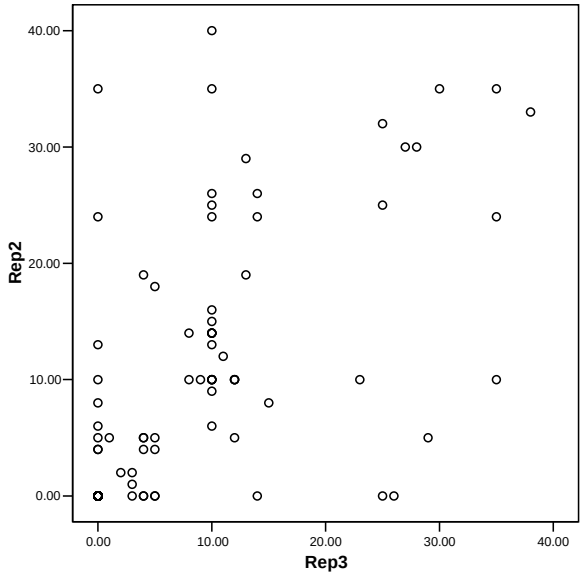


Figure 38: Comparison of the 2nd and 3rd Repetitions of Loudness in Group 2; Including all Voice Samples.

CHAPTER FOUR: DISCUSSION

An understanding of how listeners perceive voice quality is important for the clinical assessment and management of patients with voice disorders. Since rating scales are commonly used to study voice quality, the primary focus of this study was to examine listener agreement and reliability of clinician's rating of perceptual parameters utilizing the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) protocol on a sample of pediatric voices. This study included participants with differing levels of experience in the field of speech-language pathology. Group 1 consisted of students with graduate level training in the area of voice disorders, and group 2 consisted of students who had graduate level training in the area of voice disorders who also completed two training listening and rating voice samples.

Intra-Rater Reliability and Agreement of Group 1 and 2

The mean severity scores for group 1 and group 2 indicated that the participants as a whole were found to be fairly consistent across the three repetitions, for each of the perceptual measures. The greatest difference noted was 15mm. This finding was unexpected at the onset of the study, based on previous research reporting poor intra-judge reliability and agreement for perceptual judgments of voice quality (Blaustein & Bar, 1983; DeBoer & Shealy, 1995; Kreiman et al., 1993, Kreiman & Gerratt, 1998). For example, Kreiman et al. found perceptual ratings to be so variable that multiple listeners' ratings for a given voice sample spanned the entire range of a seven-point rating scale. However, a recent study by Shrivastav and colleagues (in press) suggests that this finding is not surprising given that individual differences in the perception of voice quality may not be as large as concluded in previous studies. This argument is based on

the psychometric theory stating that part of the variability in inter and intra-listener ratings results from differences in how listeners use rating scales, rather than true differences in the perception of voice quality itself (Shrivastav, Sapienza, & Nandur, in press). Under this assumption, the authors postulate that if inter-listener variability in perception is relatively small, it may be possible to generate a model of voice quality that is applicable to the average listener.

Other observations by Wolfe, Martin, and Palmer (2000) found that beginning clinicians seem to base their judgments more on the severity of the voice sample rather than specific attributes present in the sample. The mean severity scores from group 1 were found to follow along this trend, with the perceptual indices (roughness, breathiness and strain) mirroring the scores from overall severity. Additionally, the authors stated that due to its saliency, degree of severity is not a factor that beginning clinicians need extensive training on in order to be fairly consistent. Interestingly, the mean severity scores for group 1 were fairly consistent across the three repetitions of each voice sample. In contrast, group 2 underwent the perceptual training program to distinguish the difference between vocal quality parameters and to identify the perceptual characteristics that were most prevalent in the voice sample. The mean severity scores demonstrated that the clinicians were judging the sample for more than overall severity, and were also able to distinguish between other characteristics they heard.

Although the mean severity averages were found to be fairly consistent within each group, the standard deviations for both groups were found to be quite high. The standard deviations for group 2 were not significantly lower than group 1. Kreiman, et al. (1993), discussed the levels at which clinicians can reliability distinguish perceptual differences. For example, if a beginning clinician can only distinguish five levels of breathiness on a 100 point scale then the difference between 54, 63, and 53, (in the current study for group 1, voice sample

7), may hold no meaning. For numerous samples, both groups made a mark above or half way in between the mild, moderate, or severe marker on the 100 mm scale. This may offer an explanation for the high standard deviations that were seen for group 1 and group 2.

The variability in a listener's response to the same physical stimulus across multiple trials has also been addressed by others. These studies emphasize that the listener's response to a physical stimulus results from a number of factors, only some of which are directly related to the physical characteristics of the stimulus. When the goal of a perceptual task is to determine how changes in a the physical characteristics or stimulus affect listeners' perception, extraneous factors such as those related to time, frequency of occurrence of stimuli and "internal conditions" introduce errors in perceptual scaling (Graham, 1950). Shrivastav et al. (in press) discuss two kinds of errors that are frequently identified when using a rating scale for a psychophysical task. The first kind, which may be called "random errors," are those that cannot be predicted or adequately controlled. These types of errors arise from factors such as rapid changes in attention, listener fatigue, chance, etc., and lead to random variability in a listener's response. The second kind of error is introduced by factors that have a systematic effect on all perceptual responses made by a listener within a particular time frame. For example, a listener who is more conservative in using the rating scale may use a narrow range to rate all stimuli as compared to a listener who is more liberal in the use of the same rating scale. Additionally, two listeners may differ on what they consider to be a particular point on the rating scale (for example, the term "moderate" may mean different things to different listeners). Although there is some evidence to suggest that factors such as the number and range of stimuli tested and the order of presentation of stimuli may affect perception itself (Briada & Durlach, 1972; Lacouture, 1997; Luce,

Nosofsky, Green, & Smith, 1982), these effects can be controlled to some extent by randomizing the order of presentation of stimuli (Shrivastav et al., in press).

Examining the mean severity scores for group 1 and group 2 revealed that overall the participants were reliable in the rating of severity for the six perceptual indices. Therefore, the alternative hypothesis for hypotheses 1 and 2 were accepted indicating that there is a strong correlation within each group when comparing severity scores.

Listener Agreement between Group 1 and Group 2

The mean severity scores for group 2, as compared to group 1, were consistently lower across the perceptual indices roughness, breathiness and strain. For example (Table 13 and 14), group 1 ranked voice sample one for strain as 23mm, 19mm, and 21mm, and group 2 ranked voice sample one as 6mm, 9mm, and 9mm. Group 1 ranked voice sample three for strain as 13mm, 14mm, and 16mm; whereas group 2 ranked voice sample three 6mm, 6mm, and 6mm. One explanation for this finding may be that individual clinicians differed in the parameters they considered important (Kreiman et al., 1990). ANOVA results revealed that for strain the mean scores were statistically significant for 8 out of the 12 voice samples, as compared to roughness in which only 3 out of the 12 voice sample differences were significant. One plausible explanation for this finding may be that strain is perceptually harder to isolate, and clinicians may require additional training to be able to differentiate from other parameters.

The means for the perceptual terms loudness and pitch were remarkably lower in group 2 than in group 1. Analysis of variance revealed that for loudness 8 out of the 12 voice sample were significantly different, and 7 out of the 12 voice samples were significantly different for pitch. Anecdotal reports from clinicians indicate that they felt that "even with an increase in

roughness and strain, there was little to no effect on the perceived pitch or loudness of the voice sample."

With the examination of the mean severity scores for group 1 and group it was observed that there were significant differences present among all the six perceptual indices. Therefore, the alternative hypothesis for hypotheses 3 was accepted indicating a change in the perceptual rating of clinicians with an increased level of training.

Strengths of the Study

It is obvious from this study that training of a specific task (such as perceptual ratings) shows a trend of improvement in variability within and between participants, and that these results call attention to the need of more specific training. Another strength of this study was the exposure of graduate students to a standardized scaling method for the perceptual rating of voice disorders. The participants were able to easily understand and employ the CAPE-V, because of chosen perceptual indices and the utility of the visual analogue scale. Additionally, for the carryover into a clinical practice, the CAPE-V allows for a clinician to make baseline measures and track the progress of a client without the need of expensive equipment.

Limitations of the Study

A limitation of the study was the small sample size, of twelve participants per group. A second limitation may be observed in a 'learning factor,' because the voice samples were not presented with a time lapse between the repetitions, there was a potential learning effect, or recognition of the voice samples when listened to the second and third time. This learning effect could have impacted the level of agreement that was seen for the second and third repetitions.

Future Research

To further examine the reliability of the CAPE-V with a pediatric population the voice samples should be presented to expert clinicians specializing in the evaluation and treatment of voice disorders. This would allow for further identification of the level of experience that is needed to improve consistency of perceptual measures. It would be of interest to see what perceptual characteristics the expert clinicians noted are the primary factors affecting each of the voice samples. Future studies could also involve the presentation of the voice samples to a group of trained and untrained participants with one week interval between the repetitions. This would help understand a 'learning effect' relationship that may exist when making multiple ratings of voice perception in the same data collection session. Additionally, the samples could be presented to a group of untrained participants, and then have them attend a training session and examine their rating before and after training.

Summary

For the voice clinician, there is value in the reliability of perceptual ratings. For a new clinician, in an expanding field, this is a skill that is necessary to cultivate in order to ensure success when treating a voice disordered population. This study found that beginning clinicians were able to make ratings that are reliable, and with additional training the variability of those ratings decreased. Therefore study calls for the development of a training mechanism in order to further centralize the meaning of perceptual terms, it's anatomical and physiological correlate, and understand the range of severity. With this level of understanding patients therapeutic needs may be met and/or improved within centers and across the voice community.

APPENDIX A: CAPE-V EVALUATION FORM

Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)

Name:_____ **Date:**_____

The following parameters of voice quality will be rated upon completion of the following tasks:
1. Sustained vowels, /a/ and /i/ for 3-5 seconds duration each.
2. Sentence production:
   a. The blue spot is on the key again.
   b. How hard did he hit him?
   c. We were away a year ago.
   d. We eat eggs every Easter.
   e. My mama makes lemon muffins.
   f. Peter will keep at the peak.
3. Spontaneous speech in response to: "Tell me about your voice problem." or "Tell me how your voice is functioning."

Legend: C = Consistent I = Intermittent
MI = Mildly Deviant
MO =Moderately Deviant
SE = Severely Deviant

SCORE

Overall Severity _____ C I /100
MI MO SE

Roughness _____ C I /100
MI MO SE

Breathiness _____ C I /100
MI MO SE

Strain _____ C I /100
MI MO SE

Pitch (Indicate the nature of the abnormality):
_____ C I /100
MI MO SE

Loudness (Indicate the nature of the abnormality):
_____ C I /100
MI MO SE

_____ _____ C I /100
MI MO SE

_____ _____ C I /100
MI MO SE

COMMENTS ABOUT RESONANCE: NORMAL OTHER (Provide description):

ADDITIONAL FEATURES (for example, diplophonia, fry, falsetto, asthenia, aphonia, pitch instability, tremor, wet/gurgly, or other relevant terms):

Clinician:_____

APPENDIX B: GROUP 1 DATA

Participant	Sample	OS1	OS2	OS3	Mean	SD	R1	R2	R3	Mean	SD
1	1	9	10	55	25	26	9	10	45	21	21
1	2	35	56	40	44	11	35	56	50	47	11
1	3	10	10	8	9	1	10	12	9	10	2
1	4	64	58	63	62	3	64	59	61	61	3
1	5	54	59	60	58	3	35	59	51	48	12
1	6	61	54	63	59	5	61	45	62	56	10
1	7	70	75	65	70	5	70	75	70	72	3
1	8	54	41	64	53	12	54	42	59	52	9
1	9	22	10	10	14	7	11	15	10	12	3
1	10	10	44	34	29	17	10	37	30	26	14
1	11	10	10	9	10	1	10	10	7	9	2
1	12	3	44	12	20	22	2	17	6	8	8
2	1	18	15	29	21	7	17	51	39	36	17
2	2	32	62	62	52	17	47	57	56	53	6
2	3	29	34	19	27	8	29	23	56	36	18
2	4	61	60	68	63	4	60	61	60	60	1
2	5	53	58	62	58	5	53	47	49	50	3
2	6	56	56	60	57	2	53	54	54	54	1
2	7	69	69	66	68	2	35	57	53	48	12
2	8	40	54	58	51	9	46	50	44	47	3
2	9	26	18	16	20	5	54	22	16	31	20
2	10	43	39	51	44	6	48	43	49	47	3
2	11	15	15	15	15	0	15	15	14	15	1
2	12	6	10	15	10	5	26	15	16	19	6
3	1	10	10	10	10	0	10	10	0	7	6
3	2	54	35	54	48	11	60	35	62	52	15
3	3	10	10	10	10	0	10	0	0	3	6
3	4	55	55	35	48	12	75	54	54	61	12
3	5	55	55	54	55	1	74	60	62	65	8
3	6	22	54	35	37	16	21	53	10	28	22
3	7	54	65	75	65	11	53	65	75	64	11
3	8	23	35	35	31	7	35	53	23	37	15
3	9	4	5	10	6	3	10	5	21	12	8
3	10	10	10	20	13	6	10	10	10	10	0
3	11	5	10	10	8	3	0	0	0	0	0
3	12	10	10	0	7	6	10	0	0	3	6
4	1	35	57	44	45	11	16	65	18	33	28
4	2	65	54	28	49	19	29	74	40	48	23
4	3	20	32	1	18	16	50	40	0	30	26
4	4	84	59	48	64	18	90	71	50	70	20
4	5	51	49	44	48	4	65	55	54	58	6
4	6	57	59	54	57	3	55	50	65	57	8
4	7	53	59	57	56	3	29	63	49	47	17
4	8	52	50	39	47	7	49	50	41	47	5
4	9	0	44	15	20	22	0	52	10	21	28
4	10	25	44	30	33	10	20	35	48	34	14

4	11	10	3	0	4	5	0	0	0	0	0
4	12	18	6	0	8	9	3	8	0	4	4
5	1	14	18	8	13	5	29	17	22	23	6
5	2	56	61	74	64	9	58	56	69	61	7
5	3	0	39	0	13	23	0	28	0	9	16
5	4	54	70	74	66	11	60	64	60	61	2
5	5	64	74	74	71	6	64	62	75	67	7
5	6	76	44	45	55	18	75	43	40	53	19
5	7	88	82	75	82	7	73	79	78	77	3
5	8	38	48	55	47	9	48	55	42	48	7
5	9	2	18	6	9	8	14	16	19	16	3
5	10	18	38	24	27	10	31	30	55	39	14
5	11	0	0	0	0	0	0	0	0	0	0
5	12	5	4	0	3	3	0	6	0	2	3
6	1	9	10	55	25	26	9	10	45	21	21
6	2	35	56	40	44	11	35	56	50	47	11
6	3	10	10	8	9	1	10	12	9	10	2
6	4	64	59	63	62	3	64	59	61	61	3
6	5	54	59	60	58	3	35	59	51	48	12
6	6	61	54	63	59	5	61	45	62	56	10
6	7	70	75	65	70	5	70	75	70	72	3
6	8	54	40	64	53	12	54	42	59	52	9
6	9	22	10	10	14	7	11	15	10	12	3
6	10	10	44	34	29	17	10	37	30	26	14
6	11	10	10	9	10	1	10	10	7	9	2
6	12	3	44	12	20	22	3	17	6	9	7
7	1	47	19	29	32	14	49	14	15	26	20
7	2	70	44	49	54	14	68	44	48	53	13
7	3	19	38	32	30	10	21	41	19	27	12
7	4	65	63	61	63	2	63	67	66	65	2
7	5	55	59	62	59	4	50	66	61	59	8
7	6	60	31	48	46	15	55	56	40	50	9
7	7	64	75	65	68	6	77	75	54	69	13
7	8	36	64	65	55	16	40	64	53	52	12
7	9	27	16	16	20	6	37	16	19	24	11
7	10	33	37	32	34	3	30	45	17	31	14
7	11	10	14	14	13	2	10	16	13	13	3
7	12	18	17	14	16	2	15	14	15	15	1
8	1	55	8	4	22	28	75	0	0	25	43
8	2	35	17	32	28	10	56	24	54	45	18
8	3	10	15	32	19	12	5	0	14	6	7
8	4	43	42	32	39	6	60	90	59	70	18
8	5	44	54	30	43	12	58	50	15	41	23
8	6	41	28	43	37	8	20	10	40	23	15
8	7	70	67	74	70	4	40	16	44	33	15
8	8	59	67	50	59	9	18	60	30	36	22
8	9	10	5	6	7	3	0	0	0	0	0
8	10	30	18	26	25	6	15	0	19	11	10
8	11	28	0	0	9	16	0	0	0	0	0
8	12	9	10	10	10	1	0	0	0	0	0

9	1	10	0	28	13	14	10	0	10	7	6
9	2	35	45	52	44	9	35	47	54	45	10
9	3	0	35	22	19	18	0	0	0	0	0
9	4	49	60	70	60	11	55	35	70	53	18
9	5	40	54	60	51	10	47	35	60	47	13
9	6	40	25	49	38	12	40	0	10	17	21
9	7	53	70	70	64	10	50	70	70	63	12
9	8	45	40	54	46	7	40	0	30	23	21
9	9	24	54	54	44	17	0	57	15	24	30
9	10	24	0	35	20	18	0	0	35	12	20
9	11	0	37	0	12	21	0	0	0	0	0
9	12	0	0	0	0	0	0	0	0	0	0
10	1	10	11	10	10	1	11	11	11	11	0
10	2	35	35	35	35	0	36	35	35	35	1
10	3	11	9	10	10	1	11	0	11	7	6
10	4	70	70	35	58	20	70	70	70	70	0
10	5	51	71	47	56	13	54	70	70	65	9
10	6	35	35	35	35	0	35	35	35	35	0
10	7	70	71	69	70	1	70	70	71	70	1
10	8	35	70	70	58	20	35	70	70	58	20
10	9	10	10	11	10	1	10	9	4	8	3
10	10	9	10	24	14	8	0	10	34	15	17
10	11	1	2	2	2	1	0	0	0	0	0
10	12	20	14	34	23	10	10	1	10	7	5
11	1	45	22	19	29	14	20	34	18	24	9
11	2	54	68	53	58	8	56	62	60	59	3
11	3	20	22	24	22	2	20	29	18	22	6
11	4	60	57	57	58	2	55	65	56	59	6
11	5	74	70	59	68	8	58	62	58	59	2
11	6	70	45	59	58	13	54	49	59	54	5
11	7	91	90	92	91	1	85	86	94	88	5
11	8	61	59	63	61	2	58	54	55	56	2
11	9	12	4	12	9	5	4	8	21	11	9
11	10	21	45	18	28	15	26	46	26	33	12
11	11	1	3	1	2	1	1	0	2	1	1
11	12	5	18	17	13	7	4	27	22	18	12
12	1	3	25	5	11	12	3	10	0	4	5
12	2	50	54	55	53	3	14	54	55	41	23
12	3	0	35	10	15	18	0	1	0	0	1
12	4	56	35	56	49	12	24	25	10	20	8
12	5	35	70	54	53	18	10	35	35	27	14
12	6	61	56	55	57	3	10	10	10	10	0
12	7	70	9	96	58	45	3	35	35	24	18
12	8	35	35	35	35	0	10	20	24	18	7
12	9	10	1	1	4	5	10	1	1	4	5
12	10	25	5	5	12	12	0	0	0	0	0
12	11	0	0	0	0	0	0	0	0	0	0
12	12	10	10	4	8	3	0	0	0	0	0

Participant	Sample	B1	B2	B3	Mean	SD	St1	St2	St3	Mean	SD
1	1	9	10	45	21	25	9	10	55	25	30
1	2	10	55	34	33	41	35	49	34	39	41
1	3	10	49	16	25	30	10	48	15	24	29
1	4	62	49	56	56	54	53	58	59	57	58
1	5	65	51	40	52	48	53	54	48	52	51
1	6	63	55	54	57	55	62	45	54	54	51
1	7	70	75	71	72	73	69	75	71	72	73
1	8	55	41	59	52	51	35	44	64	48	52
1	9	0	35	10	15	20	0	10	10	7	9
1	10	15	35	34	28	32	21	42	34	32	36
1	11	10	18	5	11	11	35	2	5	14	7
1	12	3	17	6	9	11	2	16	5	8	10
2	1	29	50	38	39	42	14	55	25	31	37
2	2	42	60	54	52	55	57	57	54	56	56
2	3	43	45	50	46	47	27	24	33	28	28
2	4	59	59	56	58	58	58	59	56	58	58
2	5	44	54	50	49	51	49	54	49	51	51
2	6	49	50	63	54	56	50	55	61	55	57
2	7	52	53	48	51	51	57	65	53	58	59
2	8	44	55	49	49	51	54	63	50	56	56
2	9	48	25	18	30	24	25	20	18	21	20
2	10	46	36	44	42	41	45	30	49	41	40
2	11	15	15	15	15	15	15	15	14	15	15
2	12	10	15	18	14	16	9	14	18	14	15
3	1	21	23	10	18	17	9	10	0	6	5
3	2	53	53	63	56	57	53	44	35	44	41
3	3	10	22	34	22	26	10	0	0	3	1
3	4	75	63	55	64	61	53	54	35	47	45
3	5	75	75	62	71	69	53	52	35	47	45
3	6	35	61	54	50	55	35	61	22	39	41
3	7	75	75	75	75	75	75	75	75	75	75
3	8	35	60	54	50	55	10	20	20	17	19
3	9	0	10	20	10	13	0	0	0	0	0
3	10	35	10	35	27	24	0	0	0	0	0
3	11	0	10	4	5	6	0	0	0	0	0
3	12	10	10	0	7	6	0	0	0	0	0
4	1	45	53	38	45	45	21	34	40	32	35
4	2	60	50	34	48	44	74	59	38	57	51
4	3	61	44	6	37	29	78	43	0	40	28
4	4	71	54	63	63	60	80	60	54	65	60
4	5	49	59	44	51	51	32	63	60	52	58
4	6	39	73	54	55	61	55	50	80	62	64
4	7	34	59	44	46	50	67	66	51	61	59
4	8	59	58	48	55	54	53	39	38	43	40
4	9	0	31	18	16	22	0	39	14	18	24
4	10	33	33	52	39	41	6	48	34	29	37

4	11	3	0	0	1	0	0	6	0	2	3
4	12	7	13	0	7	7	15	15	0	10	8
5	1	9	16	14	13	14	14	13	5	11	10
5	2	20	50	64	45	53	40	64	65	56	62
5	3	5	30	0	12	14	0	37	0	12	16
5	4	30	70	54	51	58	39	73	68	60	67
5	5	68	66	50	61	59	60	73	76	70	73
5	6	30	38	64	44	49	74	56	48	59	54
5	7	74	39	54	56	50	85	84	79	83	82
5	8	64	35	57	52	48	67	57	56	60	58
5	9	14	0	0	5	2	14	17	0	10	9
5	10	17	17	29	21	22	37	41	31	36	36
5	11	0	0	0	0	0	0	0	0	0	0
5	12	0	14	21	12	16	0	4	4	3	4
6	1	9	10	45	21	25	9	10	55	25	30
6	2	10	55	34	33	41	35	49	54	46	50
6	3	10	49	16	25	30	10	48	15	24	29
6	4	61	49	56	55	53	53	58	59	57	58
6	5	65	51	40	52	48	53	54	48	52	51
6	6	63	55	54	57	55	62	45	54	54	51
6	7	70	75	71	72	73	69	75	70	71	72
6	8	55	41	59	52	51	35	44	64	48	52
6	9	0	35	10	15	20	0	10	10	7	9
6	10	15	35	34	28	32	21	42	34	32	36
6	11	10	18	5	11	11	35	2	5	14	7
6	12	3	7	6	5	6	2	16	5	8	10
7	1	19	32	36	29	32	56	26	29	37	31
7	2	46	28	38	37	34	69	45	49	54	49
7	3	18	40	35	31	35	10	30	31	24	28
7	4	59	54	65	59	59	70	68	64	67	66
7	5	48	52	51	50	51	55	65	58	59	61
7	6	59	26	39	41	35	63	38	39	47	41
7	7	56	75	64	65	68	75	75	53	68	65
7	8	58	64	67	63	65	39	64	61	55	60
7	9	28	27	16	24	22	21	16	18	18	17
7	10	40	21	38	33	31	40	21	15	25	20
7	11	10	15	13	13	14	11	15	14	13	14
7	12	14	16	14	15	15	17	16	13	15	15
8	1	45	0	0	15	5	38	12	15	22	16
8	2	50	40	59	50	50	59	40	39	46	42
8	3	0	0	52	17	23	0	0	58	19	26
8	4	52	38	52	47	46	75	33	39	49	40
8	5	43	59	39	47	48	39	61	39	46	49
8	6	50	40	50	47	46	39	46	54	46	49
8	7	67	88	93	83	88	98	99	100	99	99
8	8	39	89	82	70	80	100	100	87	96	94
8	9	0	0	0	0	0	0	0	0	0	0
8	10	40	39	21	33	31	31	44	25	33	34
8	11	0	0	0	0	0	0	0	0	0	0

8	12	0	0	0	0	0	0	0	0	0	0
9	1	0	0	0	0	0	10	0	0	3	1
9	2	0	0	0	0	0	35	0	53	29	27
9	3	0	0	29	10	13	0	35	0	12	16
9	4	0	0	0	0	0	53	64	70	62	65
9	5	0	0	1	0	0	56	55	52	54	54
9	6	0	0	25	8	11	54	0	64	39	34
9	7	0	0	0	0	0	62	70	70	67	69
9	8	44	10	49	34	31	0	44	47	30	40
9	9	0	0	0	0	0	0	0	0	0	0
9	10	10	0	0	3	1	35	0	0	12	4
9	11	0	15	0	5	7	0	33	0	11	15
9	12	0	0	0	0	0	0	0	0	0	0
10	1	10	35	10	18	21	10	35	11	19	22
10	2	35	36	35	35	35	35	70	36	47	51
10	3	10	35	35	27	32	10	9	10	10	10
10	4	70	70	35	58	54	70	70	70	70	70
10	5	57	70	35	54	53	34	71	70	58	66
10	6	35	35	35	35	35	10	35	35	27	32
10	7	69	71	70	70	70	71	70	70	70	70
10	8	35	70	70	58	66	35	70	70	58	66
10	9	0	10	4	5	6	0	9	4	4	6
10	10	10	11	34	18	21	9	10	31	17	19
10	11	0	0	0	0	0	0	0	0	0	0
10	12	35	35	35	35	35	35	35	11	27	24
11	1	70	40	28	46	38	10	16	23	16	18
11	2	55	54	57	55	55	42	66	63	57	62
11	3	42	45	39	42	42	15	17	19	17	18
11	4	64	64	54	61	60	54	66	56	59	60
11	5	70	74	44	63	60	39	70	64	58	64
11	6	78	54	56	63	58	44	46	61	50	52
11	7	14	70	3	29	34	88	88	94	90	91
11	8	35	52	64	50	55	63	54	61	59	58
11	9	24	0	11	12	8	0	26	24	17	22
11	10	30	55	35	40	43	18	59	14	30	34
11	11	2	4	2	3	3	1	0	0	0	0
11	12	1	43	26	23	31	1	19	20	13	17
12	1	19	35	10	21	22	0	10	4	5	6
12	2	40	54	55	50	53	70	34	55	53	47
12	3	18	35	10	21	22	10	35	10	18	21
12	4	35	55	57	49	54	35	23	35	31	30
12	5	20	54	35	36	42	60	70	70	67	69
12	6	63	55	54	57	55	61	70	54	62	62
12	7	70	70	37	59	55	70	97	97	88	94
12	8	35	35	35	35	35	54	35	35	41	37
12	9	10	1	1	4	2	10	1	10	7	6
12	10	35	10	10	18	13	10	0	0	3	1
12	11	0	0	0	0	0	0	0	0	0	0
12	12	10	10	4	8	7	0	0	0	0	0

Participant	Sample	P1	P2	P3	Mean	SD	L1	L2	L3	Mean	SD
1	1	34	10	55	33	33	9	10	35	18	21
1	2	10	40	34	28	34	10	30	40	27	32
1	3	10	10	8	9	9	10	5	6	7	6
1	4	58	49	60	56	55	51	35	42	43	40
1	5	35	54	53	47	51	35	54	55	48	52
1	6	45	50	54	50	51	44	31	41	39	37
1	7	65	64	71	67	67	64	64	55	61	60
1	8	35	34	68	46	49	34	34	66	45	48
1	9	0	10	10	7	9	1	10	10	7	9
1	10	10	17	10	12	13	10	10	19	13	14
1	11	35	4	10	16	10	10	10	10	10	10
1	12	3	9	6	6	7	3	9	10	7	9
2	1	15	44	34	31	36	16	47	39	34	40
2	2	44	49	50	48	49	37	41	49	42	44
2	3	19	14	20	18	17	17	13	18	16	16
2	4	62	52	47	54	51	54	48	48	50	49
2	5	53	53	43	50	49	40	52	42	45	46
2	6	43	44	49	45	46	49	47	49	48	48
2	7	44	48	34	42	41	34	54	34	41	43
2	8	44	49	39	44	44	43	48	40	44	44
2	9	9	22	17	16	18	9	18	16	14	16
2	10	34	40	44	39	41	39	40	45	41	42
2	11	15	15	14	15	15	14	15	15	15	15
2	12	12	15	17	15	16	11	15	16	14	15
3	1	0	0	4	1	2	0	0	5	2	2
3	2	34	22	35	30	29	34	10	20	21	17
3	3	10	0	0	3	1	10	0	0	3	1
3	4	10	35	10	18	21	10	10	10	10	10
3	5	10	10	0	7	6	10	10	10	10	10
3	6	0	10	10	7	9	10	22	10	14	15
3	7	10	23	23	19	22	35	24	35	31	30
3	8	0	10	10	7	9	10	10	10	10	10
3	9	0	0	0	0	0	0	0	0	0	0
3	10	10	0	4	5	3	10	0	10	7	6
3	11	0	0	0	0	0	10	10	10	10	10
3	12	0	0	0	0	0	22	20	4	15	13
4	1	13	57	50	40	49	29	34	19	27	27
4	2	53	39	14	35	29	79	8	5	31	15
4	3	24	49	0	24	24	9	15	0	8	8
4	4	68	57	38	54	50	40	37	49	42	43
4	5	32	47	34	38	40	25	29	20	25	25
4	6	40	41	35	39	38	62	52	8	41	34
4	7	12	42	35	30	36	29	50	20	33	34
4	8	40	41	26	36	34	29	26	15	23	21
4	9	0	50	9	20	26	0	29	8	12	16

4	10	14	49	25	29	34	5	39	18	21	26
4	11	6	2	0	3	2	45	5	0	17	7
4	12	16	8	0	8	5	54	13	0	22	12
5	1	4	0	5	3	3	0	0	0	0	0
5	2	0	39	26	22	29	40	15	16	24	18
5	3	0	0	0	0	0	0	0	0	0	0
5	4	8	13	29	17	20	0	0	1	0	0
5	5	31	38	27	32	32	40	3	0	14	6
5	6	39	18	20	26	21	40	12	0	17	10
5	7	73	30	26	43	33	43	15	0	19	11
5	8	28	18	18	21	19	8	0	0	3	1
5	9	0	0	0	0	0	0	0	0	0	0
5	10	0	16	25	14	18	0	0	0	0	0
5	11	0	0	0	0	0	0	0	0	0	0
5	12	17	0	0	6	2	15	0	0	5	2
6	1	34	10	55	33	33	9	10	35	18	21
6	2	10	40	34	28	34	10	30	40	27	32
6	3	10	0	8	6	5	10	5	6	7	6
6	4	58	49	60	56	55	51	35	42	43	40
6	5	35	54	53	47	51	35	54	55	48	52
6	6	45	50	54	50	51	44	33	41	39	38
6	7	65	64	70	66	67	64	64	55	61	60
6	8	35	34	68	46	49	34	34	66	45	48
6	9	0	10	10	7	9	1	10	10	7	9
6	10	10	17	27	18	21	10	19	18	16	18
6	11	35	4	10	16	10	10	10	10	10	10
6	12	3	9	6	6	7	3	9	10	7	9
7	1	24	19	29	24	24	21	13	29	21	21
7	2	69	41	49	53	48	36	40	50	42	44
7	3	18	31	31	27	30	17	18	35	23	25
7	4	69	65	64	66	65	46	35	35	39	36
7	5	60	63	62	62	62	54	54	54	54	54
7	6	63	29	54	49	44	60	58	56	58	57
7	7	66	75	68	70	71	66	75	69	70	71
7	8	39	64	62	55	60	36	64	63	54	60
7	9	20	15	17	17	16	20	17	17	18	17
7	10	39	41	43	41	42	29	31	20	27	26
7	11	20	15	14	16	15	10	15	13	13	14
7	12	15	15	17	16	16	29	13	26	23	21
8	1	0	4	4	3	4	0	0	0	0	0
8	2	0	20	1	7	9	39	0	0	13	4
8	3	0	45	29	25	33	20	0	0	7	2
8	4	23	5	30	19	18	57	0	1	19	7
8	5	45	0	30	25	18	24	55	0	26	27
8	6	59	34	24	39	32	58	23	38	40	34
8	7	47	75	85	69	76	80	61	54	65	60
8	8	40	44	44	43	44	55	43	47	48	46
8	9	0	0	0	0	0	53	47	27	42	39
8	10	19	15	39	24	26	28	0	24	17	14

8	11	46	0	0	15	5	94	0	24	39	21
8	12	15	18	0	11	10	5	0	28	11	13
9	1	10	0	0	3	1	10	0	35	15	17
9	2	35	0	0	12	4	2	25	0	9	11
9	3	0	0	0	0	0	0	0	0	0	0
9	4	0	0	0	0	0	10	44	58	37	46
9	5	0	35	0	12	16	4	0	0	1	0
9	6	0	40	0	13	18	0	35	0	12	16
9	7	10	0	0	3	1	0	0	0	0	0
9	8	0	0	0	0	0	0	0	0	0	0
9	9	0	0	18	6	8	35	54	10	33	32
9	10	0	0	0	0	0	0	0	10	3	4
9	11	0	35	0	12	16	0	10	0	3	4
9	12	0	0	0	0	0	0	0	0	0	0
10	1	10	10	2	7	6	0	10	2	4	5
10	2	10	10	10	10	10	10	9	36	18	21
10	3	1	9	10	7	9	0	4	2	2	3
10	4	35	35	36	35	35	35	35	10	27	24
10	5	35	35	35	35	35	35	35	35	35	35
10	6	9	35	35	26	32	10	10	10	10	10
10	7	71	36	35	47	39	35	35	35	35	35
10	8	10	34	35	26	32	10	10	37	19	22
10	9	10	8	7	8	8	35	35	37	36	36
10	10	10	35	9	18	21	0	10	10	7	9
10	11	0	0	0	0	0	0	0	0	0	0
10	12	10	1	36	16	18	0	0	35	12	16
11	1	9	10	18	12	13	19	20	20	20	20
11	2	49	41	49	46	45	53	35	52	47	45
11	3	15	19	11	15	15	10	18	15	14	16
11	4	43	54	42	46	47	56	39	45	47	44
11	5	43	56	45	48	50	60	54	55	56	55
11	6	43	35	46	41	41	54	46	43	48	46
11	7	57	76	65	66	69	65	62	77	68	69
11	8	55	44	62	54	53	54	59	60	58	59
11	9	2	0	0	1	0	2	0	0	1	0
11	10	19	25	19	21	22	29	37	20	29	29
11	11	2	2	0	1	1	2	1	0	1	1
11	12	1	20	11	11	14	14	42	30	29	34
12	1	1	0	0	0	0	0	0	0	0	0
12	2	0	0	0	0	0	0	0	0	0	0
12	3	0	0	0	0	0	0	0	0	0	0
12	4	0	0	0	0	0	0	0	0	0	0
12	5	0	0	0	0	0	0	10	0	3	4
12	6	0	0	0	0	0	0	9	10	6	8
12	7	3	0	0	1	0	4	0	35	13	16
12	8	0	0	0	0	0	2	0	0	1	0
12	9	0	1	1	1	1	2	1	2	2	2
12	10	0	0	0	0	0	0	0	0	0	0
12	11	0	0	0	0	0	0	0	0	0	0

12	12	0	0	0	0	0	10	10	10	10	10

APPENDIX C: GROUP 2 DATA

Participant	Sample	OS1	OS	OS3	Mean	SD	R1	R2	R3	Mean	SD
1	1	40	40	40	40	0	19	26	27	24	4
1	2	45	42	33	40	6	60	40	45	48	10
1	3	29	29	24	27	3	22	23	17	21	3
1	4	53	61	53	56	5	54	60	45	53	8
1	5	48	41	49	46	4	53	42	53	49	6
1	6	57	42	34	44	12	60	44	38	47	11
1	7	63	60	63	62	2	74	57	67	66	9
1	8	47	50	52	50	3	45	44	50	46	3
1	9	22	35	14	24	11	16	24	22	21	4
1	10	35	39	27	34	6	29	45	40	38	8
1	11	18	15	7	13	6	15	14	12	14	2
1	12	23	17	15	18	4	18	17	7	14	6
2	1	10	10	10	10	0	2	17	1	7	9
2	2	22	21	18	20	2	22	29	24	25	4
2	3	6	1	10	6	5	1	1	3	2	1
2	4	35	48	54	46	10	35	52	55	47	11
2	5	45	35	65	48	15	42	52	60	51	9
2	6	35	49	25	36	12	29	50	30	36	12
2	7	81	88	89	86	4	64	75	88	76	12
2	8	44	35	52	44	9	35	24	46	35	11
2	9	10	10	10	10	0	4	12	2	6	5
2	10	30	26	23	26	4	35	31	54	40	12
2	11	5	2	2	3	2	0	2	2	1	1
2	12	5	5	16	9	6	5	5	0	3	3
6	1	15	11	18	15	4	22	7	15	15	8
6	2	34	31	29	31	3	25	25	35	28	6
6	3	4	2	0	2	2	0	0	0	0	0
6	4	30	39	56	42	13	30	40	50	40	10
6	5	25	45	54	41	15	22	55	42	40	17
6	6	32	35	32	33	2	30	30	35	32	3
6	7	60	65	70	65	5	35	15	49	33	17
6	8	30	22	35	29	7	25	34	29	29	5
6	9	6	8	5	6	2	4	4	0	3	2
6	10	10	15	17	14	4	0	10	14	8	7
6	11	5	0	0	2	3	0	0	0	0	0
6	12	4	0	0	1	2	8	0	0	3	5
8	1	10	5	5	7	3	0	0	0	0	0
8	2	32	24	24	27	5	28	24	25	26	2
8	3	0	0	5	2	3	0	0	0	0	0
8	4	31	35	14	27	11	31	30	10	24	12
8	5	59	55	35	50	13	0	35	29	21	19
8	6	10	25	10	15	9	0	25	5	10	13
8	7	70	70	70	70	0	0	0	0	0	0
8	8	22	24	22	23	1	0	10	10	7	6

8	9	0	0	0	0	0	0	0	0	0	0
8	10	10	10	5	8	3	10	10	4	8	3
8	11	0	0	0	0	0	0	0	0	0	0
8	12	10	3	3	5	4	0	0	0	0	0
9	1	2	10	3	5	4	2	5	3	3	2
9	2	40	45	16	34	16	22	22	18	21	2
9	3	0	5	12	6	6	0	1	9	3	5
9	4	45	49	29	41	11	22	40	19	27	11
9	5	25	42	25	31	10	19	29	23	24	5
9	6	10	30	44	28	17	15	39	43	32	15
9	7	83	87	87	86	2	47	78	80	68	19
9	8	25	30	29	28	3	20	22	27	23	4
9	9	30	21	17	23	7	15	15	5	12	6
9	10	27	25	40	31	8	22	17	30	23	7
9	11	0	0	0	0	0	0	0	0	0	0
9	12	10	15	10	12	3	5	0	2	2	3
10	1	3	14	10	9	6	0	4	0	1	2
10	2	25	23	35	28	6	29	27	40	32	7
10	3	4	10	14	9	5	0	0	0	0	0
10	4	35	23	25	28	6	23	29	27	26	3
10	5	45	35	21	34	12	39	40	27	35	7
10	6	35	15	25	25	10	0	7	11	6	6
10	7	48	43	50	47	4	40	40	54	45	8
10	8	17	29	22	23	6	8	9	4	7	3
10	9	5	5	18	9	8	5	5	24	11	11
10	10	15	10	19	15	5	0	5	0	2	3
10	11	0	0	0	0	0	0	0	0	0	0
10	12	5	10	10	8	3	0	0	0	0	0
11	1	30	25	30	28	3	4	10	8	7	3
11	2	44	35	35	38	5	21	22	26	23	3
11	3	10	24	25	20	8	0	3	9	4	5
11	4	52	47	49	49	3	0	35	23	19	18
11	5	55	59	55	56	2	15	18	20	18	3
11	6	70	54	52	59	10	45	15	21	27	16
11	7	70	73	74	72	2	45	29	19	31	13
11	8	60	45	49	51	8	0	17	15	11	9
11	9	35	10	14	20	13	10	5	16	10	6
11	10	26	27	26	26	1	8	22	19	16	7
11	11	0	4	8	4	4	6	5	9	7	2
11	12	20	18	25	21	4	17	16	16	16	1
12	1	10	18	10	13	5	8	21	10	13	7
12	2	35	35	50	40	9	35	49	55	46	10
12	3	10	15	10	12	3	10	15	10	12	3
12	4	39	62	49	50	12	48	60	53	54	6
12	5	49	59	52	53	5	49	59	55	54	5
12	6	40	41	43	41	2	50	31	54	45	12
12	7	70	74	74	73	2	60	63	70	64	5
12	8	54	44	59	52	8	55	35	55	48	12
12	9	19	15	14	16	3	10	15	10	12	3

12	10	18	14	14	15	2	15	20	20	18	3
12	11	10	7	10	9	2	14	16	24	18	5
12	12	8	9	10	9	1	6	9	10	8	2
15	1	5	35	43	28	20	20	23	20	21	2
15	2	3	13	43	20	21	33	23	24	27	6
15	3	14	20	19	18	3	19	25	22	22	3
15	4	5	55	59	40	30	26	24	28	26	2
15	5	2	52	60	38	31	33	21	36	30	8
15	6	8	39	53	33	23	30	26	21	26	5
15	7	15	50	65	43	26	23	22	35	27	7
15	8	15	50	64	43	25	25	23	24	24	1
15	9	22	17	19	19	3	27	18	50	32	17
15	10	14	54	54	41	23	23	24	35	27	7
15	11	12	16	11	13	3	13	14	14	14	1
15	12	35	10	15	20	13	14	10	14	13	2
18	1	15	6	11	11	5	15	4	2	7	7
18	2	54	39	47	47	8	18	43	50	37	17
18	3	11	24	25	20	8	4	4	2	3	1
18	4	79	57	50	62	15	64	42	43	50	12
18	5	79	65	46	63	17	78	58	44	60	17
18	6	75	45	60	60	15	75	30	58	54	23
18	7	95	93	94	94	1	90	90	93	91	2
18	8	49	56	45	50	6	30	56	32	39	14
18	9	14	5	10	10	5	9	4	4	6	3
18	10	24	16	24	21	5	0	15	24	13	12
18	11	2	3	8	4	3	2	3	5	3	2
18	12	24	4	16	15	10	5	0	5	3	3
19	1	23	14	15	17	5	0	22	0	7	13
19	2	40	35	27	34	7	40	43	30	38	7
19	3	10	9	8	9	1	0	0	0	0	0
19	4	44	40	32	39	6	39	39	22	33	10
19	5	47	35	30	37	9	18	15	22	18	4
19	6	38	30	29	32	5	8	30	14	17	11
19	7	49	40	42	44	5	41	0	15	19	21
19	8	27	42	31	33	8	4	44	13	20	21
19	9	11	7	4	7	4	19	9	5	11	7
19	10	27	10	15	17	9	5	13	10	9	4
19	11	0	0	0	0	0	0	0	0	0	0
19	12	15	4	5	8	6	8	0	0	3	5
23	1	39	35	51	42	8	35	45	57	46	11
23	2	58	49	44	50	7	49	45	43	46	3
23	3	20	22	26	23	3	10	19	35	21	13
23	4	59	54	64	59	5	42	43	55	47	7
23	5	50	65	70	62	10	44	43	57	48	8
23	6	57	40	64	54	12	54	33	55	47	12
23	7	63	70	72	68	5	44	54	70	56	13
23	8	55	55	61	57	3	51	35	30	39	11
23	9	25	35	41	34	8	32	29	50	37	11
23	10	36	55	47	46	10	60	49	37	49	12

23	11	5	5	6	5	1	19	18	20	19	1
23	12	16	20	11	16	5	16	21	23	20	4

Participant	Sample	B1	B2	B3	Mean	SD	St1	St2	St3	Mean	SD
1	1	73	41	54	56	16	21	26	21	23	3
1	2	61	48	38	49	12	40	49	55	48	8
1	3	44	35	42	40	5	15	17	8	13	5
1	4	65	55	55	58	6	52	65	55	57	7
1	5	42	49	49	47	4	56	47	60	54	7
1	6	59	45	53	52	7	72	40	40	51	18
1	7	44	64	74	61	15	73	71	65	70	4
1	8	65	50	60	58	8	35	50	58	48	12
1	9	32	50	30	37	11	15	10	9	11	3
1	10	39	38	50	42	7	22	32	27	27	5
1	11	19	15	7	14	6	12	14	10	12	2
1	12	25	28	26	26	2	12	12	13	12	1
2	1	5	6	3	5	2	6	3	0	3	3
2	2	10	14	2	9	6	24	37	16	26	11
2	3	9	2	3	5	4	0	3	3	2	2
2	4	22	44	45	37	13	31	48	55	45	12
2	5	17	9	35	20	13	54	45	70	56	13
2	6	18	48	35	34	15	29	40	35	35	6
2	7	10	10	15	12	3	90	85	89	88	3
2	8	53	10	54	39	25	45	50	48	48	3
2	9	2	0	2	1	1	2	0	0	1	1
2	10	23	24	8	18	9	24	35	14	24	11
2	11	0	2	3	2	2	1	2	3	2	1
2	12	4	5	12	7	4	4	5	12	7	4
6	1	17	22	13	17	5	0	0	0	0	0
6	2	40	40	30	37	6	0	0	15	5	9
6	3	7	7	0	5	4	0	0	0	0	0
6	4	40	35	63	46	15	0	15	32	16	16
6	5	39	39	54	44	9	10	30	39	26	15
6	6	38	35	14	29	13	3	0	4	2	2
6	7	18	20	52	30	19	65	65	83	71	10
6	8	39	40	40	40	1	0	7	15	7	8
6	9	0	22	0	7	13	0	0	0	0	0
6	10	21	23	30	25	5	0	0	0	0	0
6	11	7	0	0	2	4	0	0	0	0	0
6	12	0	0	0	0	0	0	0	0	0	0
8	1	10	5	8	8	3	0	0	0	0	0
8	2	33	24	25	27	5	20	0	0	7	12
8	3	0	0	10	3	6	0	0	0	0	0
8	4	35	35	15	28	12	22	20	0	14	12
8	5	60	60	35	52	14	52	35	35	41	10
8	6	24	25	10	20	8	0	0	0	0	0
8	7	52	35	55	47	11	70	70	70	70	0
8	8	22	29	30	27	4	23	23	0	15	13

8	9	0	0	0	0	0	0	0	0	0	0
8	10	0	10	10	7	6	0	0	0	0	0
8	11	0	0	0	0	0	0	0	0	0	0
8	12	10	5	5	7	3	0	0	0	0	0
9	1	3	0	3	2	2	2	5	0	2	3
9	2	5	9	15	10	5	39	44	5	29	21
9	3	0	2	5	2	3	0	0	4	1	2
9	4	5	15	12	11	5	35	22	5	21	15
9	5	0	15	9	8	8	14	45	14	24	18
9	6	4	22	20	15	10	4	15	27	15	12
9	7	19	20	20	20	1	83	85	88	85	3
9	8	5	18	5	9	8	12	18	5	12	7
9	9	5	5	15	8	6	5	18	8	10	7
9	10	15	7	15	12	5	14	18	20	17	3
9	11	0	0	0	0	0	0	0	0	0	0
9	12	14	15	12	14	2	0	0	1	0	1
10	1	4	13	10	9	5	0	0	0	0	0
10	2	22	22	35	26	8	10	5	25	13	10
10	3	3	10	20	11	9	0	0	0	0	0
10	4	40	20	30	30	10	14	19	8	14	6
10	5	32	30	22	28	5	45	25	21	30	13
10	6	45	15	26	29	15	17	15	27	20	6
10	7	4	10	9	8	3	50	54	53	52	2
10	8	20	27	24	24	4	3	22	17	14	10
10	9	0	5	13	6	7	4	10	24	13	10
10	10	23	14	24	20	6	0	0	18	6	10
10	11	0	0	0	0	0	0	0	0	0	0
10	12	7	10	18	12	6	0	0	0	0	0
11	1	20	10	15	15	5	4	10	10	8	3
11	2	39	35	35	36	2	15	29	35	26	10
11	3	0	8	10	6	5	0	4	10	5	5
11	4	50	40	53	48	7	40	39	45	41	3
11	5	56	59	54	56	3	52	54	48	51	3
11	6	70	55	52	59	10	55	25	45	42	15
11	7	60	45	54	53	8	70	74	74	73	2
11	8	60	49	44	51	8	58	25	39	41	17
11	9	0	0	0	0	0	41	10	5	19	20
11	10	26	16	24	22	5	0	10	8	6	5
11	11	0	0	0	0	0	4	0	0	1	2
11	12	5	2	14	7	6	10	13	8	10	3
12	1	18	18	10	15	5	5	14	10	10	5
12	2	35	40	52	42	9	10	35	49	31	20
12	3	10	15	10	12	3	10	15	10	12	3
12	4	48	58	53	53	5	33	47	52	44	10
12	5	57	59	54	57	3	35	58	54	49	12
12	6	49	50	54	51	3	20	39	30	30	10
12	7	50	55	75	60	13	62	70	83	72	11
12	8	55	56	63	58	4	54	50	59	54	5
12	9	14	14	10	13	2	27	14	10	17	9

12	10	14	14	10	13	2	10	14	10	11	2
12	11	5	7	10	7	3	4	8	10	7	3
12	12	6	7	10	8	2	5	6	70	27	37
15	1	35	49	50	45	8	10	13	10	11	2
15	2	15	40	48	34	17	12	13	15	13	2
15	3	15	15	38	23	13	13	15	12	13	2
15	4	48	52	56	52	4	10	10	18	13	5
15	5	34	51	60	48	13	12	15	14	14	2
15	6	35	48	52	45	9	10	13	11	11	2
15	7	13	14	62	30	28	38	53	62	51	12
15	8	50	54	65	56	8	22	10	47	26	19
15	9	26	15	35	25	10	24	17	19	20	4
15	10	35	55	55	48	12	10	12	12	11	1
15	11	14	15	11	13	2	12	13	12	12	1
15	12	14	10	14	13	2	13	10	14	12	2
18	1	14	0	18	11	9	0	3	2	2	2
18	2	10	38	5	18	18	49	35	59	48	12
18	3	15	28	22	22	7	0	0	2	1	1
18	4	74	45	46	55	16	76	59	42	59	17
18	5	3	47	0	17	26	7	59	48	38	27
18	6	80	38	55	58	21	74	35	57	55	20
18	7	35	86	79	67	28	95	87	95	92	5
18	8	45	54	50	50	5	44	55	12	37	22
18	9	0	1	15	5	8	5	0	3	3	3
18	10	15	3	13	10	6	0	4	20	8	11
18	11	0	0	0	0	0	2	3	5	3	2
18	12	30	10	19	20	10	5	0	3	3	3
19	1	30	13	15	19	9	0	5	3	3	3
19	2	15	8	23	15	8	40	30	29	33	6
19	3	14	15	9	13	3	10	0	5	5	5
19	4	49	15	29	31	17	50	40	30	40	10
19	5	30	21	15	22	8	60	47	35	47	13
19	6	14	15	0	10	8	37	30	36	34	4
19	7	35	8	4	16	17	64	65	65	65	1
19	8	35	33	24	31	6	15	41	40	32	15
19	9	3	5	2	3	2	0	5	0	2	3
19	10	39	6	7	17	19	10	5	20	12	8
19	11	0	0	0	0	0	0	0	0	0	0
19	12	17	7	2	9	8	5	0	0	2	3
23	1	46	47	54	49	4	26	23	49	33	14
23	2	58	61	47	55	7	35	54	50	46	10
23	3	44	29	24	32	10	22	14	21	19	4
23	4	70	64	64	66	3	53	71	73	66	11
23	5	64	64	71	66	4	45	80	74	66	19
23	6	62	60	66	63	3	56	54	66	59	6
23	7	44	62	61	56	10	66	75	90	77	12
23	8	61	76	86	74	13	58	61	70	63	6
23	9	16	43	22	27	14	28	28	45	34	10
23	10	56	60	44	53	8	19	58	36	38	20

23	11	4	2	1	2	2	3	1	1	2	1
23	12	29	0	3	11	16	2	0	2	1	1

Participan	Sample	P1	P2	P3	Mean	SD	L1	L2	L3	Mean	SD
1	1	16	19	19	18	2	13	26	14	18	7
1	2	35	34	25	31	6	25	24	14	21	6
1	3	26	22	14	21	6	29	18	5	17	12
1	4	42	45	33	40	6	20	35	10	22	13
1	5	40	35	29	35	6	20	25	10	18	8
1	6	44	33	27	35	9	27	29	13	23	9
1	7	64	49	62	58	8	39	35	30	35	5
1	8	30	34	28	31	3	28	40	10	26	15
1	9	14	13	6	11	4	15	19	4	13	8
1	10	37	37	17	30	12	15	26	10	17	8
1	11	13	13	9	12	2	15	10	8	11	4
1	12	14	13	7	11	4	13	14	8	12	3
2	1	10	1	1	4	5	5	0	0	2	3
2	2	10	0	2	4	5	15	1	3	6	8
2	3	0	3	3	2	2	0	2	3	2	2
2	4	0	10	5	5	5	35	10	10	18	14
2	5	0	4	0	1	2	0	8	0	3	5
2	6	0	0	18	6	10	23	24	10	19	8
2	7	0	4	4	3	2	0	4	4	3	2
2	8	15	10	10	12	3	4	10	35	16	16
2	9	7	0	1	3	4	30	0	14	15	15
2	10	4	0	2	2	2	22	5	1	9	11
2	11	0	2	2	1	1	14	2	2	6	7
2	12	0	5	5	3	3	5	5	29	13	14
6	1	0	0	0	0	0	0	0	0	0	0
6	2	0	0	0	0	0	0	0	0	0	0
6	3	0	0	0	0	0	0	0	0	0	0
6	4	0	0	0	0	0	0	0	0	0	0
6	5	0	0	13	4	8	0	0	0	0	0
6	6	0	0	8	3	5	0	0	0	0	0
6	7	0	0	63	21	36	0	0	0	0	0
6	8	0	0	0	0	0	0	0	0	0	0
6	9	15	15	5	12	6	0	0	0	0	0
6	10	14	0	0	5	8	0	0	0	0	0
6	11	18	0	0	6	10	0	0	0	0	0
6	12	0	0	0	0	0	0	0	0	0	0
8	1	0	0	0	0	0	0	0	0	0	0
8	2	0	0	0	0	0	0	0	0	0	0
8	3	0	0	0	0	0	0	0	0	0	0
8	4	0	0	0	0	0	0	0	0	0	0
8	5	0	0	0	0	0	23	0	0	8	13
8	6	0	0	0	0	0	10	0	0	3	6
8	7	0	0	0	0	0	0	0	0	0	0
8	8	0	0	0	0	0	0	0	0	0	0

8	9	0	0	0	0	0	0	0	0	0	0
8	10	0	0	0	0	0	0	0	0	0	0
8	11	0	0	0	0	0	0	0	0	0	0
8	12	0	0	0	0	0	0	0	0	0	0
9	1	3	0	0	1	2	3	0	0	1	2
9	2	0	0	0	0	0	0	0	0	0	0
9	3	0	0	0	0	0	0	0	0	0	0
9	4	0	0	0	0	0	0	0	0	0	0
9	5	0	0	0	0	0	0	0	0	0	0
9	6	0	0	0	0	0	0	0	0	0	0
9	7	0	0	0	0	0	0	0	0	0	0
9	8	0	0	0	0	0	0	0	0	0	0
9	9	0	0	0	0	0	0	0	0	0	0
9	10	0	0	0	0	0	0	0	0	0	0
9	11	0	0	0	0	0	0	0	0	0	0
9	12	0	0	0	0	0	0	0	0	0	0
10	1	0	0	0	0	0	0	0	0	0	0
10	2	0	0	0	0	0	7	5	5	6	1
10	3	0	0	0	0	0	0	0	0	0	0
10	4	0	0	0	0	0	0	0	4	1	2
10	5	0	0	0	0	0	0	0	5	2	3
10	6	0	0	4	1	2	3	4	5	4	1
10	7	15	18	22	18	4	4	0	0	1	2
10	8	0	0	0	0	0	4	5	4	4	1
10	9	0	4	0	1	2	0	0	5	2	3
10	10	3	4	5	4	1	4	0	0	1	2
10	11	0	0	0	0	0	0	0	0	0	0
10	12	0	0	0	0	0	0	0	0	0	0
11	1	0	0	0	0	0	0	0	0	0	0
11	2	0	0	15	5	9	0	0	0	0	0
11	3	0	0	0	0	0	0	0	0	0	0
11	4	10	0	0	3	6	0	0	0	0	0
11	5	10	0	0	3	6	0	0	0	0	0
11	6	0	0	0	0	0	0	0	0	0	0
11	7	0	0	0	0	0	0	0	0	0	0
11	8	0	0	0	0	0	0	0	0	0	0
11	9	0	0	0	0	0	0	0	0	0	0
11	10	17	27	25	23	5	0	0	0	0	0
11	11	0	0	0	0	0	0	0	0	0	0
11	12	0	0	0	0	0	0	0	0	0	0
12	1	10	14	10	11	2	5	14	10	10	5
12	2	10	10	35	18	14	20	10	10	13	6
12	3	10	15	10	12	3	10	15	10	12	3
12	4	10	14	10	11	2	25	30	27	27	3
12	5	10	23	13	15	7	27	30	28	28	2
12	6	9	14	10	11	3	47	24	35	35	12
12	7	25	35	35	32	6	42	10	23	25	16
12	8	35	16	10	20	13	10	35	35	27	14
12	9	10	14	10	11	2	35	25	25	28	6

12	10	12	23	15	17	6	17	14	10	14	4
12	11	5	5	10	7	3	6	6	10	7	2
12	12	5	8	10	8	3	5	9	10	8	3
15	1	14	18	18	17	2	10	10	10	10	0
15	2	35	20	16	24	10	13	12	11	12	1
15	3	15	14	12	14	2	15	14	10	13	3
15	4	14	10	13	12	2	14	10	12	12	2
15	5	20	19	10	16	6	22	16	10	16	6
15	6	14	14	11	13	2	14	14	10	13	2
15	7	13	21	12	15	5	15	10	12	12	3
15	8	15	22	14	17	4	15	19	13	16	3
15	9	15	20	14	16	3	10	10	12	11	1
15	10	19	29	20	23	6	13	13	10	12	2
15	11	13	10	11	11	2	21	10	9	13	7
15	12	15	10	10	12	3	13	10	10	11	2
18	1	3	0	0	1	2	3	0	0	1	2
18	2	0	0	0	0	0	10	0	0	3	6
18	3	0	0	0	0	0	0	0	0	0	0
18	4	0	0	0	0	0	10	13	0	8	7
18	5	0	0	0	0	0	0	10	0	3	6
18	6	0	0	10	3	6	0	0	0	0	0
18	7	0	0	0	0	0	0	0	0	0	0
18	8	0	5	0	2	3	10	5	12	9	4
18	9	9	10	0	6	6	16	8	15	13	4
18	10	35	10	15	20	13	0	0	0	0	0
18	11	0	0	0	0	0	2	0	0	1	1
18	12	0	0	1	0	1	0	6	0	2	3
19	1	0	0	0	0	0	0	0	0	0	0
19	2	0	0	0	0	0	1	4	0	2	2
19	3	10	15	6	10	5	0	0	0	0	0
19	4	0	0	0	0	0	3	0	0	1	2
19	5	0	0	0	0	0	0	5	4	3	3
19	6	0	0	0	0	0	0	0	4	1	2
19	7	14	10	8	11	3	0	0	0	0	0
19	8	2	9	4	5	4	12	5	0	6	6
19	9	0	0	0	0	0	0	0	0	0	0
19	10	0	0	0	0	0	0	0	0	0	0
19	11	0	0	0	0	0	10	0	0	3	6
19	12	0	0	0	0	0	0	4	0	1	2
23	1	13	32	35	27	12	0	0	0	0	0
23	2	20	35	35	30	9	26	24	0	17	14
23	3	0	0	0	0	0	0	0	0	0	0
23	4	12	18	39	23	14	53	33	38	41	10
23	5	18	0	25	14	13	17	0	0	6	10
23	6	20	26	29	25	5	24	35	0	20	18
23	7	53	54	50	52	2	0	0	3	1	2
23	8	12	0	35	16	18	13	0	25	13	13
23	9	19	0	25	15	13	51	0	26	26	26
23	10	24	39	43	35	10	10	32	25	22	11

23	11	2	0	0	1	1	1	0	0	0	1
23	12	9	0	0	3	5	10	0	0	3	6

APPENDIX D: MEAN AND STANDARD DEVIATION COMPARASION OF GROUP 1 AND GROUP 2

VS 1	2	3	4	5	6	7	8	9	10	11	12
Overall Severity											
Group 1 Means											
22	46	12	60	53	53	69	44	14	22	8	9
17	48	24	57	61	45	67	50	16	28	9	16
25	47	15	55	56	51	72	54	14	28	5	10
Group 2 Means											
17	36	10	42	44	39	64	37	17	23	5	15
19	33	13	45	49	37	67	40	14	25	4	10
21	33	15	48	47	39	71	43	14	26	4	11
Group 1 SD											
18	13	9	11	10	16	12	12	9	11	8	7
14	14	12	10	8	12	20	12	16	18	10	14
19	13	11	14	11	10	11	12	13	11	6	10
Group 2 SD											
13	15	8	18	19	23	20	16	11	9	6	9
12	11	10	12	12	11	17	12	11	16	6	7
16	11	9	16	16	16	15	15	10	14	4	7
Roughness											
Group 1 Means											
22	44	14	62	50	45	55	41	13	17	4	6
19	50	16	60	55	38	64	47	18	24	4	9
19	53	11	56	53	41	64	44	12	29	4	6
Group 2 Means											
10	32	6	35	34	33	47	25	13	17	6	9
15	33	8	41	39	30	44	30	12	21	6	7
12	35	9	36	39	32	53	28	16	24	7	6
Group 1 SD											
21	16	15	15	17	20	24	15	16	15	6	8
20	14	16	17	11	20	20	19	18	19	7	9
17	9	16	16	16	21	17	15	7	16	5	8
Group 2 SD											
11	13	8	16	21	24	23	10	9	19	7	6
13	10	10	17	16	12	31	9	9	14	7	8
17	12	11	11	15	18	30	18	18	16	9	8
Breathiness											
Group 1 Means											
24	44	19	53	50	47	54	47	10	27	4	8
25	50	33	52	55	45	63	51	15	25	8	15
23	53	27	50	41	50	53	58	9	31	4	11
Group 2 Means											
23	29	13	46	36	42	32	43	8	25	4	13
19	32	14	40	42	38	36	41	13	23	3	8
21	30	16	43	38	36	47	45	12	24	3	11

Group 1 SD

20	9	19	21	22	21	25	11	16	13	5	10
18	20	17	18	20	19	23	20	15	16	8	13
17	17	17	17	14	11	28	12	8	14	5	12

Group 2 SD

21	19	15	20	21	24	19	19	11	15	6	10
18	16	11	17	19	16	26	19	17	19	6	8
20	16	13	18	22	21	27	22	12	17	4	7

Strain

Group 1 Means

17	50	15	58	49	51	74	45	6	23	8	7
19	48	27	57	62	46	78	53	12	28	6	11
22	48	16	55	56	52	74	54	9	22	3	6

Group 2 Means

6	25	6	35	37	31	69	31	13	9	3	5
9	28	6	38	45	26	71	34	9	16	3	3
9	29	6	35	43	32	76	34	10	15	3	10

Group 1 SD

15	15	21	13	10	17	11	26	9	14	13	11
15	19	19	15	8	17	11	20	12	21	10	11
20	11	18	13	13	15	17	17	8	16	5	7

Group 2 SD

9	19	13	21	20	27	16	21	14	9	4	5
9	20	8	21	17	17	11	19	9	17	5	5
14	14	7	23	20	20	13	23	13	10	4	20

Pitch

Group 1 Means

13	26	9	36	32	32	44	27	3	14	13	8
14	28	15	35	37	32	45	31	10	21	6	8
22	25	10	35	32	31	43	36	7	20	4	8

Group 2 Means

6	9	5	7	5	7	15	9	7	14	4	10
7	8	6	8	7	7	16	8	6	14	3	4
7	11	4	8	7	10	21	8	5	12	3	3

Group 1 SD

12	14	8	27	19	23	27	19	6	13	17	7
18	17	18	24	22	26	28	20	15	17	11	8
22	20	11	22	21	21	29	26	7	16	6	11

Group 2 SD

6	14	8	12	13	13	22	12	7	13	7	20
11	14	8	13	12	12	20	11	8	16	5	6
11	14	5	14	11	10	25	12	8	13	4	5

Loudness

Group 1 Means

9	29	9	34	30	36	43	26	13	13	16	14
12	20	7	27	34	32	42	27	18	16	6	11

18	26	7	28	27	26	39	34	11	16	7	14

Group 2 Means

3	10	4	13	9	12	8	8	13	7	6	3
4	7	4	11	8	11	5	10	5	8	2	4
2	4	2	8	5	6	6	11	8	5	2	4

Group 1 SD

10	24	7	23	19	24	25	19	28	14	27	15
15	15	7	18	23	17	26	23	19	17	6	12
16	21	11	22	25	21	24	27	11	12	8	13

Group 2 SD

4	10	9	17	11	15	16	8	17	8	7	4
8	9	7	14	11	14	10	14	9	11	4	5
5	5	4	12	8	10	10	14	10	8	4	5

APPENDIX E: LARYNGEAL DIAGNOSIS CORRESPONGING TO THE DESIGNATED VOICE SAMPLES

Laryngeal Diagnosis Corresponding to the Designated Voice Samples

Sample	Age	Diagnosis
1	8	Vocal Nodules
2	9 1/2	Vocal Nodules
3	8	Vocal Fold swelling, reflux, nodules
4	9	Reflux, with Vocal Fold edema
5	7	Nodules, with Severe Edema
6	7	Edema with LPR
7	3	Vocal Nodules
8	11	Pre-nodular Swelling
9	13	Posterior Swelling due to Allergy
10	13	Reflux, Intercordial cyst

APPENDIX F: TRAINING PROTOCOL

Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)

ORIGIN

The CAPE-V was developed from a consensus meeting sponsored by the American Speech- Language-Hearing Association's (ASHA) Division 3: Voice and Voice Disorders, and the Department of Communication Science and Disorders, University of Pittsburgh, held in Pittsburgh on June 10-11, 2002. Attending this meeting were speech-language pathologists (SLPs) who specialize in voice disorders and invited experts in human perception (see appendix). The participants' charge was to develop standardized guidelines for auditoryperceptual evaluation of voice, based on theory and data in psychoacoustics, psychometric scaling, and voice perception. Clinical practicality and brevity of administration were also considered in developing these guidelines.

Task #1: Sustained vowel /a/.

Task #2: Sentences: Six sentences were designed to elicit various laryngeal behaviors and clinical signs. The first sentence provides production of every vowel sound in the English language, the second sentence emphasizes easy onset with the /h/, the third sentence is all voiced, the fourth sentence elicits hard glottal attack, the fifth sentence incorporates nasal sounds, and the final sentence is weighted with voiceless plosive sounds. The clinician should give the person being evaluated flash cards, which progressively show the target sentences (see below) one at a time. The sentences are: (a) The blue spot is on the key again; (b) How hard did he hit him? (c) We were away a year ago; (d) We eat eggs every Easter; (e) My mama makes lemon jam; (f) Peter will keep at the peak.

OVERALL SEVERITY: Global, integrated impression of voice deviance.

1. Roughness: Perceived irregularity in the voicing source.
-This could be caused by several types of glottal area function, and the production of this quality is more complex than a breathy voice quality.
-One cause of harshness in the vocal sample is the presence of vocal nodules.
-More than one million school age children have been diagnosed with vocal nodules, and in the pediatric patient they are typically the same color as the true vocal folds and appear on the on the free margin of the vocal fold at the junction of the anterior and middle third (Sapienza, Hoffman-Ruddy, & Baker; 2004).
2. Breathiness: Audible air escape in the voice.
-Breathy voice characteristics are found in the following voice disorders: nodules, polyps, recurrent laryngeal nerve damage, edema, postpolypectomy dysphonia (no lesion), polypoid laryngitis plus nodule, or a contact granuloma.
3. Strain: Perception of excessive vocal effort (hyperfunction).
-Laryngeal hyperfunction in children is caused by an increased tension in the extrinsic laryngeal musculature, with adduction of the false vocal folds and compression of the arytenoid cartilages. -Overtime, hyperfunction can cause swelling or edema in the true vocal folds.
Pitch: Perceptual correlate of fundamental frequency. This scale rates whether the individual's pitch deviates from normal for that person's gender, age, and referent culture.
-The perception of pitch is related to the size of the laryngeal structures. The length of newborns vocal folds are on average 2.5-3.0 mm and continually grow until around twenty years of age.
-Around ten to fourteen years, there is the appearance of sex differences in vocal fold length (Sapienza et al., 2004).
-The average fundamental frequency for an adult male is 126 Hz, with a standard deviation of 26.52 Hz.
-For the adult female, the average is 181 Hz with a standard deviation of 45.18 Hz (Xue & Fucci; 2000). ---For a child, the average fundamental frequency observed is 250 Hz (Zajac, Farkas, Dindzans, & Stool; 1993).
Loudness: Perceptual correlate of sound intensity. This scale indicates whether the individual's loudness deviates from normal for that person's gender, age, and referent culture.
-The severity level of the loudness scale indicates whether there is deviation from normal based upon the patient's gender, age and referent culture (ASHA Division 3, 2004).

LIST OF REFERENCES

ANSI. (1960). *USA Standard: Acoustical terminology (SI.1).* New York: American National Standards Institute, Inc.

Apple Computer, Inc., (2005). Final Cut Pro 3.1. Cupertino, CA.

ASHA Division 3 (2002), Consensus Auditory Perception of Voice. Retrieved July 16, 2004, from http://www.asha.org/about/membership/divs/div_3.html

Bassich, C., & Ludlow, C. (1986). The use of perceptual methods by new clinicians for assessing voice quality. *Journal of Speech and Hearing Disorders*, 51, 125- 133.

Blaustein, S., & Bar A. (1983). Reliability of perceptual voice assessment. *Journal of Communication Disorders*, 16, 157-161.

Briada, L. D., & Durlach, N. I. (1972). Intensity perception. II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America*, 51(2(Part 2)), 483-502.

Carding, P., Carlson, E., Epstein, R., Mathieson, L., & Shewell, C. (2000). Formal perceptual evaluation of voice quality in the United Kingdom. *Log Phon Voc* ,25, 133-138.

De Bodt, M., Wuyts, F., Van de Heyning, P., & Croux, C. (1997). Test-retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, 11, 74-80.

DeBoer, K., & Shealy, R. (1995). The effect of voice lessons on the clinical and perceptual skills of graduate students in speech-language pathology. *Journal of Voice*, 9, 118-126.

Doehring, D. G. (1996). *Research strategies in human communication disorders.* Austin: Pro-Ed.

Fairbanks, G. (1960). *Voice and articulation drill book*. New York: Harper & Row.

Graham, C. H. (1950). Behavioral, perception and the psychophysical methods. *Psychological Review*, 57, 108-118.

Hirano, M. (1981). *Clinical examination of voice.* New York: Springer-Verlag.

Hufnagle, J. (1982). Acoustic analysis of fundamental frequencies of voices of children with and without vocal nodules. *Perceptual and Motor Skills*, 55, 427-432.

Kreiman J., Gerratt B.R. (1998). Validity of rating scale measures of voice quality. *Journal of the Acoustical Society of America*, 104 (3 Pt 1), 1589-1608.

Kreiman J., Gerratt B.R. (2000). Sources of listener disagreement in voice quality assessment. *Journal of the Acoustical Society of America,* 108, 1867-1876.

Kreiman, J., Gerratt, B., Kempster, G., Erman, A., & Berke G. (1993). Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *Journal of Speech Hearing Research,* 39, 31-40.

Kreiman J., Gerratt B.R., Precoda K. (1990). Listener experience and perception of voice quality. *Journal of Speech Hearing Research,* 33, 103-115.

Lacouture, Y. (1997). Bow, range, and sequential effects in absolute identification: a response-time analysis. *Psychology Research,* 60(3), 121-133.

Luce, R. D., Nosofsky, R. M, Green, D. M. & Smith, D. M. (1982). The bow and sequential effect in absolute identification. *Perception and Psychophysics*, 32, 397-408.

Maxwell, D. L., & Statke, E. (1997). *Research and statistical methods in communication disorders*. Baltimore: Williams and Wilkins.

Sapienza, C., Hoffman-Ruddy, B., & Baker, S. (2004). Laryngeal structure and function in the pediatric larynx: Clinical applications. *Language, Speech, and Hearing Services in Schools*, 35, 299-307.

Schiavetti, N. & Metz, D. E. (2002). *Evaluating research in communicative disorders*. Boston: Allyn and Bacon.

SPSS, Inc., (2003). SPSS 12.0 for Windows. Chicago.

Van der Torn, M., Verdonck-de Leeuw, I., Kuik, D., & Mahieu, H. (2002). Communicative suitability of voice following radiotherapy for T1 glotic carcinoma: Testing the reliability of a rating instrument. *Journal of Voice,* 16, 398-407.

Wirz, S.L. (1995). *Perceptual approaches to communication disorders.* London: Whurr.

Wolfe, V., Martin, D., & Palmer, C. (2000). Perception of dysphonic voice quality by native listeners. *Journal of Speech Hearing Research,* 43, 697-705.

Wolfe, V., & Steinfatt, T. (1987). Prediction of vocal severity within and across voice types. *Journal of Speech and Hearing Research,* 30, 230-240.

Xue, A., & Fucci, D. (2000). Effects of race and sex on acoustic features of voice analysis. *Perceptual and Motor Skills*, 91, 951-958.

Zajac, D., Farkas, Z., Dindzans, L., & Stool, S. (1993). Aerodynamic and laryngographic

assessment of pediatric vocal function. *Pediatric Pulmonology*, 15, 44-51.