2020-04-08

An Investigation into the Effects of Traumatic Brain Injury on Speech and Gait

Mir Haaris Sultan
Worcester Polytechnic Institute

Shikha Pandey
Worcester Polytechnic Institute

An Investigation into the Effects of Traumatic Brain Injury on Speech and Gait

An Interactive Qualifying Project Report

Submitted to the Faculty of

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the Degree of Bachelor of Science

by:

Shikha Pandey
Mir Haaris Sultan

Advisors:
Mark Claypool, Professor of Computer Science
Adam Lammert, Assistant Professor of Biomedical Engineering

Date: April 07, 2020

Abstract

Even mild traumatic brain injury (TBI) can lead to permanent damage if not treated immediately. In order to provide faster and more convenient methods to monitor TBI, speech and gait were explored as concussion indicators. Smartphones were used as sensors to record audio and acceleration data on the speech and gait of ten concussed and twelve non-concussed participants. Data from the two groups was statistically analyzed to determine significant differences with the hope that it could eventually lead to a more efficient method of diagnosing concussions. Although collecting speech and gait data through the smartphone was quick and convenient, there were no statistically significant differences found between the means, standard deviation, kurtosis, and skewness of the two groups.

i. Acknowledgements

This project would not have been possible without the assistance of the kind people at WPI and the surrounding area.

We would firstly like to thank all of the WPI undergraduate and graduate students who participated in our survey and study process.

And finally, our biggest thanks is reserved to our Project Advisors. Professor Mark Claypool and Adam Lammert, the excellent guidance, corrections, patience, and support that each of you provided us throughout this project helped to make it a success and a fantastic learning experience!

ii. Table of Contents

iii. List of Figures

iv. List of Tables

1.0 Introduction

As soon as Alex's town announced that there would be a championship game played at the school, tickets sold out within a matter of hours. Alex is the headlining player of his school's soccer team and all projections assert that he will be in a professional league one day. Practices have been grueling to ensure the team will be ready for the biggest game of the season; yet, just as the two week mark approaches, Alex slide-tackles a teammate during a particularly arduous practice and scores a nearly impossible goal but falls hard on his head. His vision momentarily blacks out and, as he regains consciousness, he begins to feel woozy. There is a good chance that he may have suffered a concussion, but the big game is nigh. Alex's health is in jeopardy and so is winning the game, but he is unsure of whether to seek medical help or keep his potential injury in the shadows.

First of all, why is it important if Alex is concussed? A concussion is a form of mild traumatic brain injury caused by an impact to the head (Finklestein et al., 2006). These types of injuries can impede brain function. The number of brain functions that are impeded and the degree of impedance can greatly vary, which makes concussions a dangerous problem (Finklestein et al., 2006). If Alex is concussed, he may experience nausea, loss of balance, and other types of losses in cognitive function (Murdoch et al., 2001). This would cause detriment to the two aspects most important in this situation: Alex's physical/mental health and his potential professional career. Alex has two weeks before the game, which is just enough time for him and his team to make adjustments, but Alex still requires the confirmation that he is concussed. What could help Alex and people in a similar situation is a concussion test that could be conducted in a quick and convenient way.

There are many different types of testing in the form of diagnostic aids that can be used to evaluate concussions. A popular test for athletes in college and highschool is the ImPACT Baseline test (Schatz, 2010). This test is meant to analyze an athlete's reaction time, memory, and processing speed (Schatz, 2010). However, this methodology of testing comes with several drawbacks. This ImPACT test only analyzes the mental aspects of concussions, whereas traumatic brain injury also leaves an impact on the physical aspects such as balance. The

ImPACT test is also obtrusive, invasive, and consumes a lot of time, which prevents it from being regularly used to track concussions. As such, physical changes may be easier to identify and more helpful in providing additional clarity when serving as a diagnostic aid for a concussion. Therefore, it was hypothesized prior to the data collection and analysis of the study, that there would be statistically significant speech and gait data differences between the concussed and non-concussed groups.

Gait and speech are two physical functions that may be directly affected by a concussion because of the large and consistent presence that they have in a typical, day-to-day life. Gait and speech are well developed by adulthood and often require distinct and finely coordinated patterns of movement (Patel et al., 2012). Therefore, measuring speech and gait can be a quick route to finding the most direct symptoms of a concussion (Cantu, 2012). Recording speech and gait before a concussion can also help to analyze the extent of traumatic brain injury (TBI), (Kontos et al., 2012). Since these functions are so developed and finely controlled, examining them for any changes post-concussion to pre-concussion may provide data leading to an improved method of concussion diagnosis. There has been a growing amount of research in academia that has been done on how speech and gait markers affect TBI assessment. Studies have shown that between 30% and 65% of people with TBI suffer from dizziness and disequilibrium (lack of balance while sitting or standing) at some point in their recovery (Peterson, 2015). Furthermore, motor speech disorders are common after TBI as vocal deviation of amplitude, perturbation, and pitch can be greatly affected (McHenry, 2009).

The goal of this project was to provide data and preliminary analysis on those data, that may help develop diagnostic aids for concussions in an efficient, non-intrusive way that considers both the mental and physical aspects of a concussion. This study was attempted to collect data in a relatively new area of research, while trying to find different yet effective ways of detecting concussions. This study aimed to use speech and gait markers, to provide information through compiled and analyzed data that could help confirm whether these two cognitive functions can be proper non-intrusive components of diagnosing TBI. Given how instrumental sports in college can be in determining the careers of the students, it is important that an issue such as TBI diagnosis be further explored in order to protect students. Therefore, students at WPI were targeted in the study as the sole demographic. A protocol was developed that utilized mobile phone apps to record speech and track the gait of healthy participants with no history of TBI and of participants that have suffered at least one concussion. Furthermore, a survey was developed to gather data on other potential causes or symptoms that are synonymous with those of concussions (see Appendix C). This practice of concussion testing removed the obtrusiveness of the testing, and the stigma against frequent testing by medical professionals due to the convenience of using a mobile app.

The remainder of the report is organized as follows. Chapter 2 provides a detailed background on the history and modern management of TBI and the special relations it has to speech and gait. Chapter 3 provides details on the preparation for collecting speech and gait data from the concussed and non-concussed participants through descriptions of the IRB application, our methods of analyzing the participant data, construction of data pipelines meant to assist with analysis, pilot studies, project promotion/advertisement, and the enacted procedure. Chapter 4 is where the project data is analyzed and it is divided into two sections: the results and the discussion. Chapter 5 describes whether this discovery meant our project goal was achieved, how our predictions fared, potential sources of error, and concludes what the data overall states about the project. Chapter 6 contains recommendations on how this project should be picked up in the future tying in personal experiences and technical and conceptual changes that would make the future of the project successful.

2.0 Background

This section discusses concussions in regards to history, modern effects, and current methods of treatment. Through compiling this type of information on concussions, this chapter specifies a problem and points to a potential solution that prior concussion research corroborates.

2.1 Concussion Diagnosis History

The definition of concussions is constantly evolving with the modernization of medicine. Previously established criteria utilized to diagnose concussions are often debated in medical literature, and as researchers and doctors work together to further advancements in the field of medicine, the definition of concussions continues to reshape and redefine itself. Today, there is currently no single definition of concussions, or even minor head injuries and mild traumatic brain injuries that is universally accepted (Sharp et al., 2015). In 2016, the National Academy of Engineering defined a concussion as "a brain injury induced by biomechanical forces, that may or may not cause loss of consciousness, and that typically causes rapid onset of short-lived impairment of neurological function with no abnormalities visible on standard structural neuroimaging studies" (Budinger, 2016). This definition is unique from older definitions as previous classification required a loss of consciousness and associated amnesia, but it was eventually discovered that this was not always the case. Physically, the brain floats in a pool of cerebrospinal fluid within the skull. When the head suffers an impact, the brain shifts or shakes within its pool inside the skull. Ultimately, this disruption damages brain cells and causes chemical and cellular changes within the brain that vary depending on the locations affected and the severity of the injury.

2.2 The Origin of Brain Trauma and Behavioral Changes

The origin of the relationship between brain trauma and behavioral changes has been an object of study for centuries now. An issue that neurologists regularly struggle with is whether a concussion is associated with lesser degrees of diffuse structural change as seen in severe traumatic brain injury, or if the mechanism is caused entirely by reversible functional changes (McCrory, 2001). This confusion began in ancient times when no distinction was made between the transient effects of concussions and severe traumatic brain injury (McCrory, 2001). The first clear separate recognition of concussions was made by the Persian physician, Rhazes, in the 10th century (McCrory, 2001). Italian physician, Lanfrancus, subsequently expanded this concept as brain "commotion" in the 13th century, although other Renaissance physicians continued to ignore this concept (McCrory et al. 2001). By the 18th century, a variety of hypotheses for concussions had emerged. The 19th century discovery of petechial hemorrhagic lesions in patients with severe traumatic brain injury led to a basis for establishing differences with other TBIs and concussions (McCrory et al. 2001). In 1848, Phineas Gage was impaled by a sharp iron spike while constructing new railway lines, but made a miraculous recovery. After the fact, however, family and friends claimed that he underwent an extreme personality shift; it was enough to make him appear a completely different person (Costandi, 2010).

Today, almost 160 years later, the same personality changes are found to accompany brain pathology in victims of multiple episodes of brain trauma (Budinger, 2016). One cohort study of 490 children who experienced a mild traumatic brain injury (mTBI) prior to age 14, who had no prior history of psychiatric illness, found that these children were significantly more likely to harbor psychiatric issues in the three years following injury than uninjured control subjects (Daneshvar et al., 2011). The children most commonly showcased attentional problems in the first year following injury. Despite this, there was no difference observed in children who had already had a prior history of psychiatric illness in the year preceding the injury (Daneshvar et al., 2011).

2.3 The Four Events That Spurred Concussion Discussion

Prior to the 1800s, concussions were unexplored and uncategorized; most people were simply not aware of them. Four distinct historical events were responsible for bringing concussions to the forefront of medical and official discussion. These four events garnered the attention of physicians, military officials, political representatives, and the public to bring about research that sought to understand the importance of concussions and their dangerous effects. The first event concerns one of the most controversial topics regarding concussions in the present day: football. The first official, recorded game of college football was played in 1869 and in the years following and up to the early 1880s, there were 18 deaths and 159 serious injuries sustained by college football players (Harrison 2014). In response, discussion was spurred to increase the safety of the sport (Harrison 2014).

The second event occurred during World War I and continued after its termination in 1918. Prior to this, during the American Civil War, psychiatrists tending to soldiers claimed that injury to the head or body had no direct correlation to psychiatric behavior or paralysis. In World War I, however, controversial diagnosis was released claiming long-lasting physical and mental injury in soldiers following exposure to artillery shell blasts (Jones, 2005).

The turning point in discussions of causation and diagnosis of concussions took place in 1980, during which time posttraumatic stress disorder (PTSD) was formally defined and recognized by the American Psychiatric Association (Budinger, 2016). The recognition of PTSD led into the third historical event, which took place during the height of conflict within the Middle East. The usage of improvised explosive devices (IEDs) in urban warfare saw a marked increase in incidents of brain trauma (Budinger, 2016). Studies of troops who deployed to Iraq and Afghanistan have found that service members who have suffered a concussion or mild traumatic brain injury are far more likely to develop PTSD (Hamilton, 2016).

The fourth event concerns football, but this time in the National Football League (NFL). NFL players who experienced behavioral changes and became depressed and suicidal after playing were examined in cognitive studies. In these studies, it was found that the pathology of their brains was similar to those who sustained concussions from heavy contact sports such as boxing and IED related injuries (McKee et al., 2014). From this research, cerebral traumatic encephalopathy (CTE) was diagnosed from the brains of NFL players postmortem (McKee et al., 2014).

2.4 Prevalence and Effects

TBI has become a very common problem in the present day. There are estimated to be at least 1 million emergency department attendances each year in the UK for example, due to head injuries, 90% of which have been considered to be mild. (Sharp et al., 2015). Even so, mild TBI is often overlooked and considered relatively harmless (Sharp et al., 2015). Long-term effects can be detrimental and many patients who have long term effects of concussions all report similar symptoms (Sharp et al., 2015). Most patients suffering a mild TBI recover in the first three months, but around one third of these patients report symptoms persisting beyond six months (Sharp et al., 2015). TBI can lead to long-term effects including epilepsy and neurodegeneration and increased risk of Alzheimer's disease, Parkinson's disease and CTE (Sharp et al., 2015). Epidemiological studies show increased mortality rates after mild TBI and one large cohort study tracked patients with TBI of all severities attending emergency departments in Glasgow, UK in 1995 and 1996 (Daneshvar et al., 2011). Thirteen years after the injury, the mortality rate of the group had reached over 40%, with increased mortality even in young patients (Daneshvar et al., 2011). Patients with mild TBI also had higher mortality rates than those with other types of injury (Sharp et al., 2015). In the United States, approximately 1.7 million people sustain a TBI annually; these injuries account for 1.4 million emergency room visits and 275,000 hospitalizations each year (Daneshvar et al., 2011).

The symptoms of a concussion may take some time to resolve, resulting in significant long-term burden (Daneshvar et al., 2011). When the symptoms of concussion persist as a variety of cognitive, somatic, and behavioral changes, these lingering deficits comprise Post-Concussion Syndrome (Daneshvar et al., 2011). The burden of concussions can also have detrimental societal and economical impacts. According to the World Health Organization (WHO), head injuries are a leading cause of the global burden of death and disability for all age groups below age 60 (Finkelstein et al., 2006). Estimating the magnitude of this burden is critical both for assessing the relative burden of injuries compared with other preventable health problems within a population, and for national investment for specific injury prevention activities. Averaged across all injuries, estimated percentage of lifetime productivity potential lost due to injury was 0.26% per injury (Finkelstein et al., 2006). Furthermore, the estimated amount of days lost due to injury was 15 days (Finkelstein et al., 2006).

The overall injury burden in the United States has exceeded $400 billion in medical care costs with falls and impact by or against a force, both commonly resulting in head injuries, accounting for more than 30% of injury costs (Finkelstein et al., 2006). It can also be noted that previously, most reported concussions were a result of falls or motor vehicle collisions

(Finkelstein, 2006). Recent studies of younger populations suggest most concussions occur during sporting events, with greatest risk during competitions (Graham et al. 2014). As concussions increasingly become a serious topic of interest in the medical world, advancements have been made to help aid in their diagnosis.

2.5 Progression of Diagnostic Tools

The four events described in Section 2.3 were able to shed light on the substantial impact that concussions can have for the first time in history. As a result, due to the severity of the symptoms that concussions present, multiple diagnostic tests were developed to identify the symptoms. These symptoms may include headaches, blurred vision, slurred speech, nausea, and trouble maintaining balance (Graham et al., 2014). The current cornerstone of concussion diagnosis is confirming the presence of a constellation of signs and symptoms(e.g. headaches, blurred vision, etc.) after an individual has experienced a hit to the head or body (Graham et al., 2014). Symptoms are self-reported by the athlete, often using a symptom scale (Graham et al., 2014). Although reliance on an athlete's self-report of symptoms is a fundamental component of concussion diagnosis, it can also be complicated and unreliable (Graham et al., 2014). Such reports have a subjective nature and the possibility of an athlete underreporting symptoms in order to remain on the field is likely (Graham et al., 2014). Traditional neuroimaging techniques, such as standard computed tomography (CT) and magnetic resonance imaging (MRI), were invented in the 1970s. Although these tests are now conducted to ensure that there is no bruising or bleeding occurring in the brain, they are not specific to concussion diagnosis (Graham et al., 2014). Rather, these techniques are leveraged only to rule out severe, life-threatening head and brain injuries such as skull fractures and intracranial hemorrhages (Graham et al. 2014).

Nowadays, a large number of concussion assessment tools are designed to be taken in a quick, efficient manner and are often conducted over the computer. But, the use of cellular devices and their ability to help aid in the understanding and diagnosis of concussions can not be overlooked. In this growingly digital world, technology has become the bridge to better understanding the medical field, and improving quality of life. The smart-phone has been developed with a variety of different tools and apps that can be utilized to measure speech and gait markers. Many apps today have the ability to record speech and measure values such as acceleration to to keep track of a concussed subject's gait consistency. The consistency, and easiness of using a smartphone to measure these specific markers, as a device that is so readily available and common in modern society, helps to identify a concussed subjects' diagnosis timeline, and progress towards recovery.

2.6 Modern Day Concussion Assessment and Diagnostic Aids

A medical diagnosis is the process of determining which disease or condition explains a person's signs or symptoms. A doctor trained in diagnosing and managing concussions, in order to diagnose patients, performs a detailed exam and cognitive test checking certain factors such as vision, reflexes, balance, memory, speech, etc. This helps understand the grade and severity of the concussion (Broglio et al., 2014). In contrast, a diagnostic aid is often used to diagnose signs of a concussion which can include a series of cognitive tests. Yet, the cognitive tests alone cannot diagnose a concussion. A neurovestibular exam must be a part of the evaluation. The doctor ultimately uses exams, diagnostic aids, and even imaging to confirm the concussion diagnosis. (Broglio et al., 2014).

One of the most famous diagnostic aids in ImPACT baseline testing, which was developed in the late 1990s when neuropsychologist Dr. Mark Lovell was conducting neurocognitive research studies at the Henry Ford Health System in Detroit (Covassin et al., 2009). ImPACT Testing, (Immediate Post-Concussion Assessment and Cognitive Testing) is a computerized measurement tool designed to categorize the extent of brain injury incurred by analyzing its effect on cognitive function, pre and post trauma. (Covassin et al., 2009). The test evaluates verbal recognition, memory, visual processing speed, learning, attentional processes, reaction time, and numerical sequencing ability (Covassin et al., 2009).

The ImPACT test is the most widely used tool in North America, yet still has some drawbacks (Schatz and Sandel, 2012). Although it comprehensively tests for the six categories listed previously, it does not test for other factors that might display inhibited cognitive functionality. A 40% false-positive rate was also acquired from the test in a 2007 study and in 2013, another study found that ImPACT classified healthy athletes as concussion subjects anywhere between 22% and 46% of the time (Robertson, 2013). Also, it was found that users who were more forthcoming and honest with their concussion symptoms might actually display normal behavior to the test and therefore decrease its sensitivity (Dessy et al., 2017). The easy-to-use and relatively common Sport Concussion Assessment Tool (SCAT), was created by the Concussion in Sports Group in 2005 to be utilized as a sideline screening method (Rădoi et al., 2019). Proctoring of the test does not require any type of formal neurophysiological knowledge or training. While the test is valuable as a rapid assessment tool, it is not meant to completely replace comprehensive neurological tests and should not be used as a standalone practice (Dessy et al., 2017). Furthermore, as useful as a speedy evaluation by a sports official without formal training is for efficiency, a false or hasty evaluation to place the player back in the game could prove to be detrimental to the player's wellbeing. Moreover, an additional impediment to identifying concussions in this fashion is that symptoms may not become apparent until several hours after injury (Duhaime et al., 2012).

Another supplemental test used as a concussion diagnostic aid is the Balance Error Scoring System (BESS), which tests for changes in balance, another major symptom of concussions. The BESS tests for balance and is performed for 20 seconds on flat and foamy surfaces in three unique stances: feet together, one leg raised, and one foot in front of the other (Bell et al., 2011). During the test, a proctor tallies the amount of errors observed (Bell et al., 2011). The test is similar to SCAT in that, although the test is useful for balance observation, it should not be used standalone as it does not take into account integral cognitive factors such as speech and gait. Furthermore, balance also worsens due to fatigue so confounding factors could affect the reliability of the test (Bell et al. 2011).

Finally, the HiMAT, High-Level Mobility Assessment Tool, is a test used to evaluate and quantify high level mobility following traumatic brain injury by enforcing tasks such as walking forward and backwards, running, jumping, hopping, ascension, and descension (Kleffelgaard et al., 2013). The HiMat has been widely used as a justification for the studying of gait markers, and has been widely regarded in prior studies done regarding gait markers for concussion study. Although the test has a high level of reliability, repeated, high-level movements can be tiring for subjects and they are typically only allotted a single trial per movement (Kleffelgaard et al., 2013).

2.7 Use of Gait Patterns

Concussions are multifaceted injuries that can affect different regions of the brain in different degrees based on the injury and the individual. As such, it can be useful to analyze multiple cognitive functions in order to aid diagnose and better treat concussions. One such methodology is to analyze locomotion and gait. Research has shown that concussed subjects display a significant change in gait patterns as concussions affect static balance, dynamic balance, static balance, vestibular–oculomotor functioning, and cognition (Howell et al., 2018). In one particular research study conducted by the University of Oregon, participants were tested against a balance and gait assessment suite (Pitt et al., 2019). It was found with statistical significance that concussed participants displayed much slower walking speeds when performing a cognitive task during the motion (Pitt et al., 2019). Moreover, it was found that concussed participants spent an inordinate amount of time on the double leg stance component of the gait cycle as compared to the single leg stance, meaning that their swing cycles were diminished (Pitt et al. 2019). In addition to this, the University of Oregon study also demonstrated a significant negative correlation in the single-leg stance and positive correlations in double-leg stance between time and the amount of concussions a participant was subjected to (King et al., 2013). The single-leg and double-leg study suggest that people who have a documented history of concussions tend to adopt conservative gait strategies to account for the compromised gait patterns as opposed to those without any history. As a result, monitoring gait is vital to the understanding of concussion symptoms.

2.8 Use of Speech Patterns

Another useful component of cognitive function is to analyze speech production. Incoherent or slurred speech and delayed processing of words, sentences, and thoughts are some examples of neurocognitive and neurobehavioral outcomes of concussions (Robbins, 2016). A growing body of evidence suggests that a concussion has the ability to not only make patients slur their words, but also elongate sounds, or speak in a more nasal tone (Robbins, 2016).

Concussions may even make a voice croak or jitter so briefly that it is imperceptible to the human ear (Robbins, 2016). In a study relating to Purdue University's Varsity Football Team, researchers addressed a population of athletes regularly receiving impacts to the head and showing signs of preclinical mild traumatic brain injury (Helfer et al., 2014). This was a state indicated by impaired cognitive performance occurring prior to a concussion (Helfer et al., 2014). In this study, results found that athletes suffered two types of changes as a result of mild TBI (Helfer et al., 2014). The first change was in average vocal tract dynamics as measured by formant frequencies, their velocities, and acceleration (Helfer et al., 2014). The second type of change was in articulatory coordination measured by a novel formant-frequency cross-correlation characterization (Helfer et al., 2014). Furthermore, researchers used these identified features to allow machine learning algorithms to detect preclinical mTBI identified by a battery of cognitive tests (Helfer et al. 2014). Murdoch, (2001), wrote an entire book surveying the many speech changes in TBI (Traumatic brain injury: Associated speech, language, and swallowing disorders). The cognitive and/or linguistic deficits were found to be associated with traumatic brain injury and multiple assessments and treatments of communicative and swallowing disorders were illustrated by a numerous amount of case studies conducted with his patients. In one study, it was also found that the detriments in auditory ability were mixed depending on the individual, with five unique subjects displaying five individualized failure methods (Turgeon et al., 2011). An alarming amount of evidence depicts the heterogeneous nature of head injuries and the unique effects they can have on people (Turgeon et al., 2011). As such, monitoring speech along with other cognitive abilities is vital to the correct understanding of concussion symptoms.

2.9 Summary

As time has progressed, society's understanding of concussions has developed in relation to breakthroughs in medical research and publicized casualties. The topic is constantly evolving and research/case studies are still being conducted in order to keep up and aid in the understanding of concussion symptoms and diagnosis. Concussions today have become an issue that is of great importance to society. The diagnosis of concussions has also progressed to encompass a wide range of tools that can assess the extent of damage without having to analyze the brain directly. Medical professionals continue to broaden the understanding of TBI through increased media presence and education on concussions and their significance. The evolution of technology, of course, introduced the smartphone, which features a plethora of useful applications, especially in the field of ease and accessibility. These facets not only enable people to make quick calls and send fast messages, but also enable researchers to speedily and efficiently track audio/speech and gait patterns. This strategy of collecting and analyzing data will serve to provide a new and improved medium when considering a test that can easily identify concussion symptoms and take into account important matters such as time efficiency, effectiveness, and the non-intrusive nature of the test itself.

3.0 Methodology

The goal of this project has been defined as providing data to assist with the creation of non-intrusive diagnostic aids to make concussion diagnosis quick and convenient. Research has established analyzing gait and speech patterns could be a viable method of achieving this goal. Therefore, the collection of speech and gait data would have to be ascertained with direct procedures and then specifically analyzed with the right equipment. This methodology details the process taken to ensure this goal was fulfilled through the collection of speech and gait data.

3.1 Overview

The primary measurements collected in this project are speech and gait. These measurements came from two distinct experimental methods. The first method ascertained a participant reading a given passage and the second method consisted of an analysis of a participant's gait as they walk 100 yards. Data was recorded through the usage of the Android apps AndroSensor [] and wav.Recorder []. Participants for the experiments were gathered in two ways. The first way was through WPI Athletics Trainers, because they had the ability to refer athletes with concussion histories. The second method of attracting participants was completed through the general student body. Emails were distributed and posters were displayed to advertise the project. Incentives for contributing data included donuts for all participants and a raffle entry to win a $25 gift card. All eligible and chosen participants were required to sign safety consent forms and complete a survey on concussion symptoms before testing.

3.2 The Procedure

Data from subjects was collected in the bottom of Perreault Lecture Hall in Fuller Labs as a great amount of participants for the study were more inclined to meet there. As a result, this became the default recording location. The survey, Androsensor, and .wavRecorder were all prepared on the cellular devices prior to data collection. Two chairs were placed on one side of a table, and one more was situated at the other side. Verbal introductions were made as soon as the participant arrived and paperwork was completed shortly thereafter. Once the recording space was finalized, data collection commenced. Participants recited "The Caterpillar Passage in a clear voice to record audio as a reference to the participants' normal speech patterns (see Appendix A). This took approximately 1-2 minutes. The participant remained seated whilst providing this data to replicate a safe and normal environment to work in comfortably. Participants were then asked to walk in order to record motion tracking data. This process involved the participants walking at their normal pace for up to 100 yards. This phase of experimentation took about 9-10 minutes. This phase only tested for a comfortable walking pace and did not place the participant in any extreme conditions.

3.2.1 Preliminary Participant Steps

The first step required the participants to sign and date two identical consent forms (see Appendix D). The participants also had to sign the forms as witnesses. Reading the forms was recommended to the participants so that they would have an idea of the risks associated with experimentation, and the extent of their rights during participation. Participants were also required to complete a survey that inquired about their sports activities, times of exercise, sleep, caffeine intake, and current and past injuries. Survey questions were attached in a document labeled "Rivermead Post Concussion Symptoms Questionnaire" (see Appendix C). Following custom, demographic questions were asked such as sex, date of birth (age), and basic physical/mental capabilities such as ease of walking up a flight of stairs or walking a mile.

3.2.2 Debriefing

All participants were thanked dearly for coming and asked for any final questions or suggestions that they may have. If asked, we provided further details regarding the project such as factors being tested, hypotheses, or discoveries made on the project. Participants were also provided with the opportunity to peruse the final paper to view analysis of data collected, should they be so inclined.

3.3 Preliminary Actions

In order for the project to be conducted, specific pieces of project equipment and school regulatory approval was required. Due to the fact that data would be collected from volunteers performing a physical activity, the first component required was written permission from the Institutional Review Board (IRB) of WPI. In order to acquire approval, a protocol describing the process of data collection was created and an application was completed and electronically dispatched to the IRB. The second preliminary action required was to obtain two smartphones, as these would be responsible for collecting the gait and speech data. Upon acquiring both the phones and IRB approval, the process of collecting data could be started.

3.3.1 Equipment Needed

This project emphasizes testing convenience so that the dangers of TBI can be quickly addressed and mitigated. Therefore, equipment that is easy to access and utilize was the only equipment operated. As previously mentioned, smartphones are an extremely useful instrument to test for concussions due to their commonality amongst athletes and trainers. Since most people possess a cellular device, an expensive instrument was not purchased for concussion testing. Furthermore, most, if not all, phones have access to the internet access, built-in audio recorders, and accelerometers. To test phones with the methodologies described earlier, two Motorola 5G smartphones were obtained that had the capability to leverage the app .wavRecorder, which records audio in a .wav format, and the app Androsensor, which collects a variety of gait associated data, including acceleration, in a .csv format.

3.3.2 WPI IRB Approval

Permission from the Institutional Review Board (IRB) of WPI is a prerequisite for working with human subjects. This approval required online training and submission of application of IRB Approval. The online training was completed through the Collaborative Institutional Training Initiative, known by the acronym CITI. CITI had a course specific to acquiring IRB approval, which taught the history and legality of human research, rights of the participants, aspects that constitute good research with human participants, and proper etiquette whilst working with human volunteers. CITI provided this online training in sections, where each section had a mandatory quiz that needed to be passed in order to move onto the next section. Once all of the sections and quizzes were completed with scores above an 80%, an electronic certificate of approval was delivered to WPI affirming that the students had proficient knowledge of the content taught within the course and was capable of collecting data from participants. The IRB also required a protocol describing the process of data collection from participants along with a paper application where academic details regarding the project were recorded. Once all these documents were submitted, feedback was received from the IRB where approval was granted and the project was allowed to continue.

3.4 Data Collection Methods

The applications being used in the experiment needed to be able to record audio in a .wav file format and provide accelerometer data. Both applications also needed to be able to export the gathered data to a desktop computer. Androsensor is an app that leverages the smartphone as a sensor, ultimately designed to analyze gait. Using Androsensor, the cellular device was held against the torso of the subject as they walked. The app simultaneously recorded gait data through three dimensional accelerometer values. Androsensor recorded the accelerometer readings and outputted these readings through a .csv file that was filed away in the phone's internal storage.

In terms of speech, there were a variety of apps on the market already designed to record audio, but for the purposes of this project, saving the audio as a .wav file was the distinction. The reason that the team required the recording to be in .wav format rather than the commonly used .mp3 format was due to the fact that .wav is the uncompressed form of the .mp3 and would prove to be more valuable in data analysis. As such, the team settled on utilizing the app .wavRecorder, as it records audio in an acceptable quality and stores the audio as a .wav file in the phone's internal storage.

It was important in the study that all speech recordings were consistent to ease the process of comparison of the subjects. Therefore, "The Caterpillar" Passage (see Appendix A) was selected as an element of consistency for all participants to read from to take part in the speech analysis. This increased the value of speech recordings, because having all participants read the same passage, would result in the data collected being more indicative of differences in motor function, and the ease with which participants read the passage would provide the team with an idea of specific speech segments to search for within the audio. Prior to the collection of speech and gait data, the participants completed a survey in regards to their current condition as well.

There are a variety of symptoms associated with a recent TBI, which can make the diagnosis extremely difficult. As such, the Rivermead Questionnaire was used in the survey to identify whether or not the participants had experienced symptoms due to TBI, or if the symptoms being experienced were due to confounding factors. In order to distinguish the participants, the initial question of the survey asked for a summary of the participant's concussion history. To dovetail back on the fact that TBI symptoms present themselves in a variety of ways, there was also an option included where the participants could report a potential concussion that they believed had never been diagnosed.

3.4.1 "The Caterpillar" Passage

"The Caterpillar" passage was developed for the primary purpose of augmenting current motor speech assessment protocols. The passage formulation itself, follows a multistep process in which researchers identify a subset of speech tasks used in motor speech evaluation, taking into consideration the nature of the sounds and syllables being spoken at that time (Patel et al., 2012). The passage has reading tasks requiring prosodic modulation, contrastive sentence types, words with emphatic stress, and multiple instances of repetition to further aid in the understanding of speech motor function (Patel et al., 2012). Additionally, DDK rates and repetition in the passage, has been used to evaluate the speech of individuals with dysarthria and speech motor disorders including those with concussions (Ben-David and Icht, 2016). Furthermore, it has been reported that non-word DDK productions may isolate more of the motor speech effects of speech disorders where real-word productions (buttercup, pattycake, caterpillar) may rely on language ability and motor speech ability to produce (Ben-David and Icht, 2016).

Due to the nature of specific instances of repetition, sounds and syllables in “The Caterpillar passage”, was chosen to represent the speech patterns of all subjects’ trials, as multi-syllable words and non-word DDK tasks are instrumental when comparing speech patterns of subjects with a history of concussions, and lack thereof (Ben-David et al., 2016).

3.4.2 Rivermead Questionnaire

A concussion diagnosis is difficult to ascertain in a quick fashion, therefore, there is a possibility that some participants may have been concussed without their knowledge. As such, the Rivermead Questionnaire was included within the survey in an effort to regulate data collected from participants. The questionnaire has been cited in over 40 research papers and is a proven resource (Ben-David and Icht, 2016). The questionnaire ultimately helped identify potential concussion symptoms so that these symptoms can be referred to in the event that the data provided shows irregularity or results that are more comparable to those expected of a concussed participant.

3.5 Pilot Studies

It was important to practice data collection methods that matched what was approved by the IRB, in order to catch any potential flaws before the experiments were conducted with consenting participants. Within this project, there were multiple tiers of data collection because both gait and speech were explored. For gait, a set method of securing the phone to the torso had to be vetted so that gait could be measured accurately. To accomplish this, there were two solutions possible: the participants could either be given an adjustable strap that could hold the phone across their waist, or they could simply hold the phone to their waist with their hand. For speech, a quiet environment with little to no ambient noise was required in order to record, so it was necessary to locate quiet locations suitable for this type of data collection. In regards to the survey, it had to be constructed so that it was easily understandable and so that the display worked efficiently on the phone’s graphical user interface (GUI). To tie the entire modus operandi together, there needed to be a script that organized the sequence of data collection and highlighted the important information to disclose, as mandated by the IRB.

3.5.1 Process

The pool of subjects managed for pilot studies was restricted just to the team and the advisors. There was a strong emphasis on practicing verbal introductions in order to satisfy IRB requirements. This included ensuring that the entire descriptional monologue flowed smoothly and concisely from signing the consent forms to asking the participant if they had any remaining questions. For gait data, it was concluded that the adjustable strap was preferable as most participants did not normally walk with their hands on their stomachs. The torso strap and manually holding the phone to the torso both yielded the same data after multiple sessions of tests on a single student. When it came to speech, there were multiple locations procured that were quiet, but did not wholly restrict occupants from conversing. Discovered locations included the WPI Library and the WPI Recreation Center. Speech recording was conducted in a conveniently selected environment, and the session consisted of the participant being seated at a table with the phone either resting in the participant's hand or on the table, recording audio. This worked well as it was coincidentally uncovered that the phone recorded higher quality audio when on the table, as opposed to in front of the participant's mouth where it caught a substantial amount of background and white noise. The procedure would also be practiced with multiple trials executed in order to test whether the time needed to conduct trials could be shortened whilst still following the IRB approved protocol. Once the data was recorded, the compilation and transfer of data would have to be coordinated as soon as possible, in order to ensure that all of the data was secured in a safe, accessible place for further analysis.

3.5.2 Observations

Through further experimentation with participants in the pilot studies, it was found that the Androsensor app could record a time increment as low as 0.005 seconds. Because this option provides the most possible data for gait, it was concluded that this lowest increment would be utilized with all the participants in the actual trials. When it came to the usage of .wavRecorder, the app functioned normally. It was shown through pilot studies that the selected applications are fully functional and store the speech and gait data in desirable formats. The survey also worked as desired because anonymity could be guaranteed and the conditional statements functioned properly in accordance with the response to questions regarding the participant's concussion history.

However, there was a major issue with all of this, despite the equipment functioning properly: the speech and gait data were not transferable from the phones. There were different exportation options provided by Androsensor and .wavRecorder, but the most useful option was to email the created files. However, when attempted, this resulted in error-ridden files being transmitted that could not be opened. Technical support personnel from the technology center in the WPI Library and Motorola support personnel were not able to provide any assistance. A solution was eventually reached when an app called Android File Transfer was discovered. This app was able to display the desired gait and speech files in a fashion that was easy to access and transferable onto the computers of the students so that they could be stored and analyzed later.

Management of the data was seen to be a much larger issue than expected, given the convenience of Androsensor and .wavRecorder. That being said, this was the only major issue experienced during the pilot studies. The script developed for interacting with participants was flexible, concise, and followed IRB rules. The physical location of audio recording was also proven to be much more flexible than initially expected. It was observed that as long as the location was relatively quiet and adequate space was available, enough to walk at least 10 meters in a straight line, there were no apparent problems with data collection.

3.6 Soliciting Participants

People with varying types of concussion histories were desired participants in order to collect and sample a wide range of data to analyze. However, it was not guaranteed that post-concussion data could be acquired from participants, or participants with concussion histories, so it was important to collect as much data as possible. WPI has classes that offer credit for participating in projects, so this was leveraged to rapidly acquire a list of participants. Participants signed themselves up for specific times on an online spreadsheet, to manage and organize participant data collection. The Social Sciences department and the Interactive Media & Game Development department both have systems where students are obligated to participate in research studies, so it was found that contacting both departments was beneficial. The primary desired source of participants was sought through a collaboration with the athletic trainers, however. WPI athletic trainers, when contacted, referred our project to groups with high concentrations of concussions in order to guarantee variation within the data. As a contingency plan, students also solicited volunteers from their own networks in order to derive a large population of users with variation in concussion history.

3.6.1 Promotion via Email and Posters

Permission was granted by WPI to place posters in high traffic areas on campus such as the Rubin Campus Center, Morgan Hall, Goat's Head & Founders Hall, Higgins Labs, Salisbury Labs, Stratton Hall, Foisie Innovation Studio, Gordon Library, and the WPI Recreation Center. Each of these locations boasts a large, diverse student presence, which means that the displayed posters would receive more publicity. The posters also advertised a $25 gift card as a potential reward, and donuts as a guaranteed reward, for participating in the survey and listed the team's contact information (see Appendix B). The goal of data collection was simplified and stated beneath a flashy hook, which was presented in such a way as to appeal to the altruistic nature of the students should they have no interest in the incentivized fried dough confections. As previously mentioned, the WPI athletic trainers were also a resource that provided a large variety in participants.

3.7 Data Pipelines

3.7.1 Gait Data Pipeline Construction

Table 1. Sample Raw Accelerometer Data recorded in m/s/s at 0.005 second increments

X (m/s¬≤)	Y (m/s¬≤)	Z (m/s¬≤)
-1.9058	2.4948	7.8434
-1.9058	2.4948	7.8434
-1.9058	2.4948	7.8434
-1.9058	2.4948	7.8434
-1.9058	2.4948	7.8434
-1.9058	2.4948	7.8434
-1.9058	2.4948	7.8434
-2.7869	3.055	11.2719
-2.7869	3.055	11.2719
-2.7869	3.055	11.2719
-2.7869	3.055	11.2719
-2.7869	3.055	11.2719

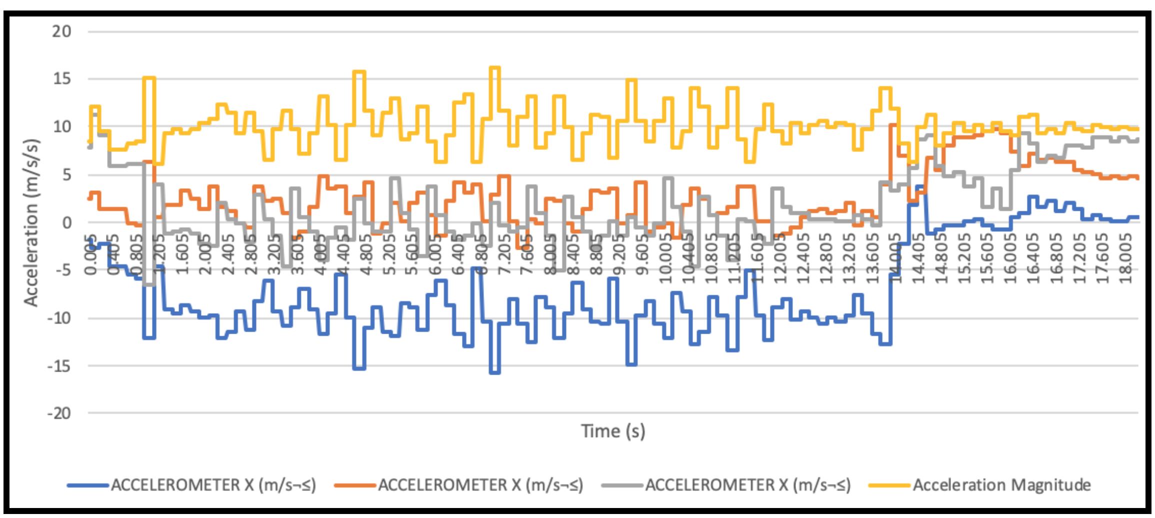


Figure 1: Acceleration in X, Y, Z orientation and acceleration magnitude plotted against time

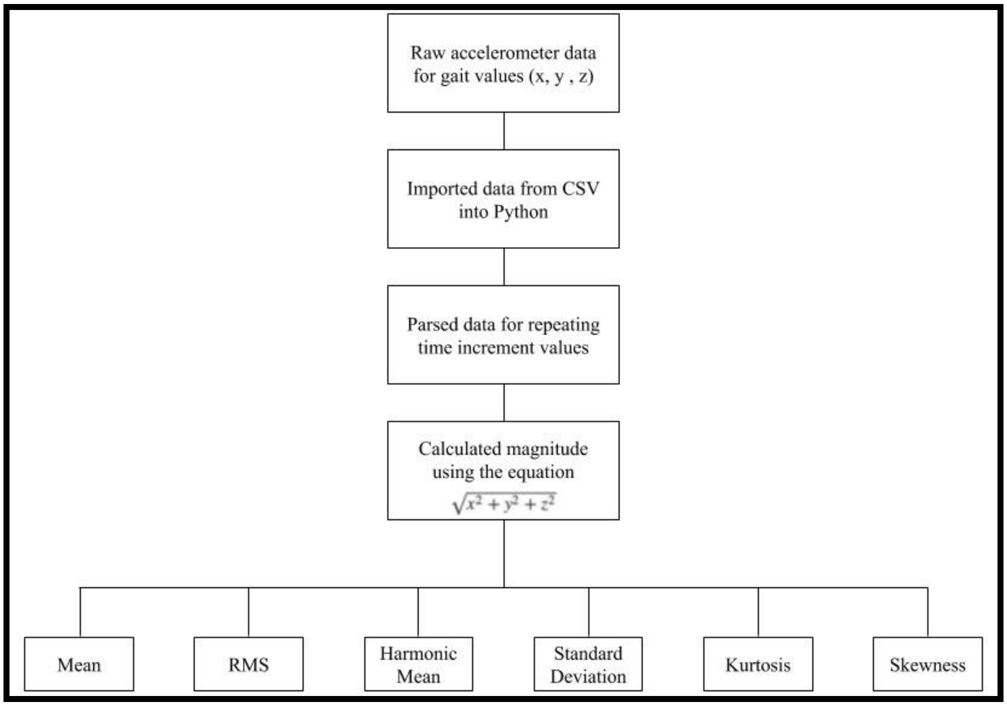


Figure 2: Diagram illustrating how gait data from .csv file is processed with Python program.

Gait data was collected from a .csv file with three accelerometer columns, corresponding to three dimensional directionality representing X, Y, and Z. As seen in Figure 1, the graph of the three directions indicates irregularities. There are no real peaks in the graph which showcases that the sensor used to record the accelerometer values could not keep up with the recording increment set at 0.005 seconds. Therefore, the sensor may have compensated by repeating previously recorded values until the sensor could readjust and record a new value. This idea is supported by the raw data seen in the .csv file in Table 1. It can be seen that there are several values repeated successively amongst the three columns for a period of time. So when it comes to analyzing this accelerometer data, it would have to be parsed before being processed for analysis. Therefore, a data pipeline was constructed in the programming language Python to clean the data and analyze the data collected. In Figure 2, an overview of how the Python pipeline was constructed is shown. The pipeline worked with one .csv file at a time where it would take the input of a .csv file and it would output the values of statistical variables derived from the .csv file.

The first step extracted the accelerometer data from the file and divided these values into three separate lists where each element was a 2-D coordinate. The x coordinate was the time in seconds, and the y coordinate was the accelerometer values in m/s/s. One list was for the X accelerometer values, one for the Y values, and one for the Z values. The reason the values were separated in this manner instead of being kept together to make it easier to process for the later steps. As seen in Table 1, the accelerometer values repeat in all three columns for a period of time. However, there are over three thousand rows of accelerometer values in the example file that Table 1 is derived from. Because that file contained the gait data of one participant, it would be time-consuming to read through all the column values to ensure the values are repeated in the same period when it can be done quickly in Python. Therefore, the second reason for separating the column values into their own variables was done so that these values could be parsed of repeating values individually and then compared with each other to see if they actually retained original values for the same time periods. Keeping the time increment linked to the individual accelerometer values allowed for the distinction to be made if the accelerometer values were recorded at the same increment. After the repeats were parsed, the three lists were compared to see if all three lists had values at the same time increments. If there was a value at a time increment that all the other lists did not have, it would be removed.

After this second step of parsing, the data was processed for analysis. The first step was to compress the three lists into one by computing the magnitude of three dimensional data, which can also be seen in Figure 1. The magnitude in the pipeline did not keep the time values as they were no longer necessary, but it was kept in Figure 1 so it could be graphed. The magnitude simplified the data analysis so that the statistical analysis could be applied to a single list of values instead of three separate ones. From there, the standard deviation, mean/average, root mean square (RMS), harmonic mean, skewness, and kurtosis were calculated from the magnitude. After those values were calculated, they were stored into a Microsoft Excel file in one of two groups: the concussed and the non-concussed. These two groups were compared to by box and whisker plots for each statistical variable so that they could be shown together to provide a visual comparison. Secondly, the groups were compared using the Data Analysis pack on Excel, where t-tests were done on all the columns of respective variable values for each group to determine if the two groups were significantly different from each other, one variable at a time. The statistical analysis applied to gait was through standard deviation, mean/average, root mean square/RMS, harmonic mean, skewness, kurtosis, and t-tests.

Standard deviation, denoted as std dev, is the measure of the amount of variation in a set of values (Petruccelli, 1999). This is important for establishing whether the values in a set can be considered similar to each other should they not deviate from the average value of the set (Petruccelli, 1999). This was useful in the gait study because it showed a measure of how much the acceleration values vary when looking for potential details to distinguish between non-concussed and concussed.

Mean, also known as the average, is the sum of observations divided by the number of observations (Petruccelli, 1999). The observations in gait were the accelerometer values, and finding the average value would be important because the average would be a fair representation of the data representation (Petruccelli, 1999). Calculating mean is important to the gait study because an average acceleration can provide a general idea of how fast the participant may have been walking which can vary from person to person.

Root mean square, or RMS, is the square root of the mean of squares (Petruccelli, 1999). This provides a measure of the magnitude of a set of values which provides a reference of the typical size of the values (Petruccelli, 1999). Because the RMS takes the magnitude, the root of the sum of all the data values squared, the calculation of RMS worked well with data measuring rates that consisted of a great amount of variation. This is directly applicable to the gait study because there is variation in the gait values due to the fact that the gait is measured in acceleration.

Harmonic Mean is one of the three Pythagorean Averages (Hayes, 2019). This value best determines the true average of a data set if the data set contains values of rates or ratios (Hayes, 2019). All the acceleration values vary as seen in Figure 1. Therefore, with the harmonic mean being particularly specialized in finding the true mean of rates, it has the potential to be especially useful for determining if there is a higher level of significance between the concussed and non-concussed groups.

Skewness is a measure in relation to the distribution curve of a data set. The skewness measures the symmetry of a data set where the value is negative if the data is skewed left and positive if the data is skewed right (Jain, 2019). This value is a measure of an entire set of data looking at the alignment of data (Jain, 2019). This is applicable to the gait study because it provides an idea of how balanced the participants are walking. Furthermore, it provides symmetry in the distribution curve of their data values providing a small skewness value.

Kurtosis is also a measure in relation to the distribution curve of a data set. Where skewness measures symmetry in the middle of the curve, kurtosis looks at the tails of the curve (Jain, 2019). If there is a higher distribution of values to the tails of the curve, the data set is said to have a high kurtosis value (Jain, 2019). This value is important because it can show whether the mean of a data set is a fair representation of the data by signifying if the data values that are considered outliers are common or rare (Jain, 2019). This applies to the gait study because it shows the significance of the data distribution and how the paces of the participants can vary.

A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features (Kenton, 2020). The t-test is one of many tests used for the purpose of hypothesis testing in statistics (Kenton 2020). To test the difference in means between the concussed and non-concussed groups, t-tests were conducted (Kenton, 2020). The aim of this is to determine whether there is any significant difference between the values calculated for each group as it is hypothesized that there will be a difference between the gait data of the non-concussed group and the concussed group.

3.7.2 Speech Data Pipeline Construction

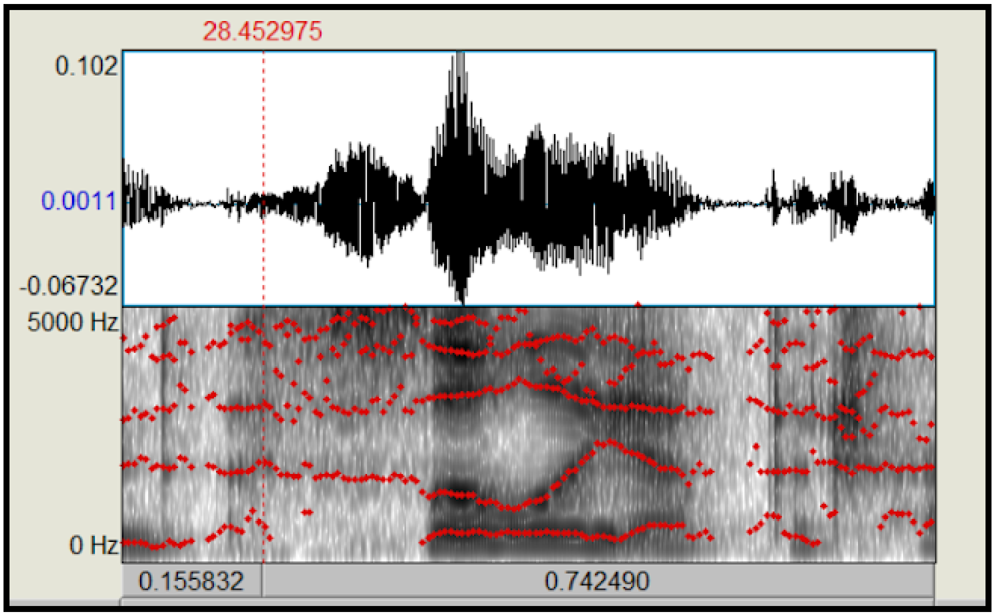


Figure 3: Waveform graph and spectrogram of example User 87.

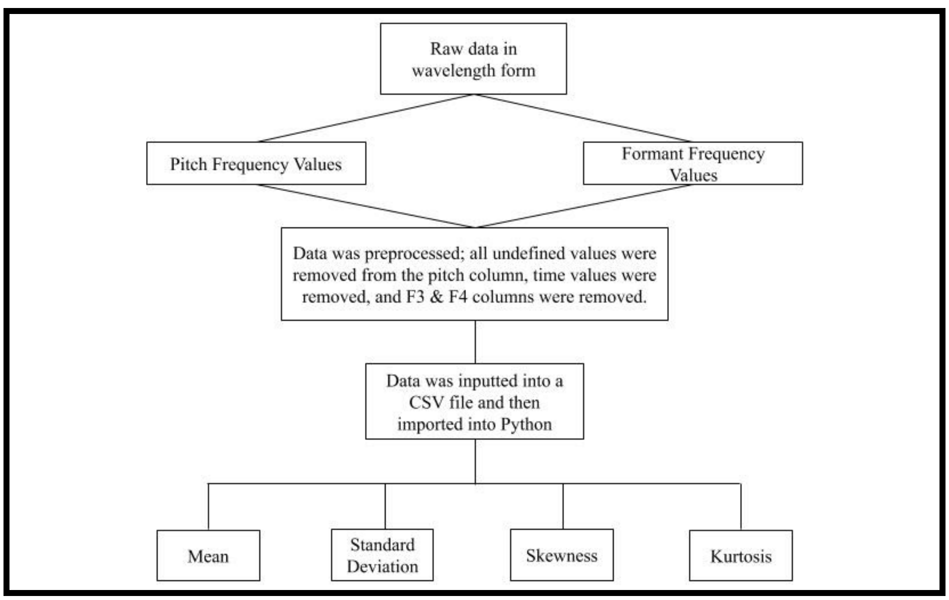


Figure 4: Diagram illustrating how speech data is processed with Python program

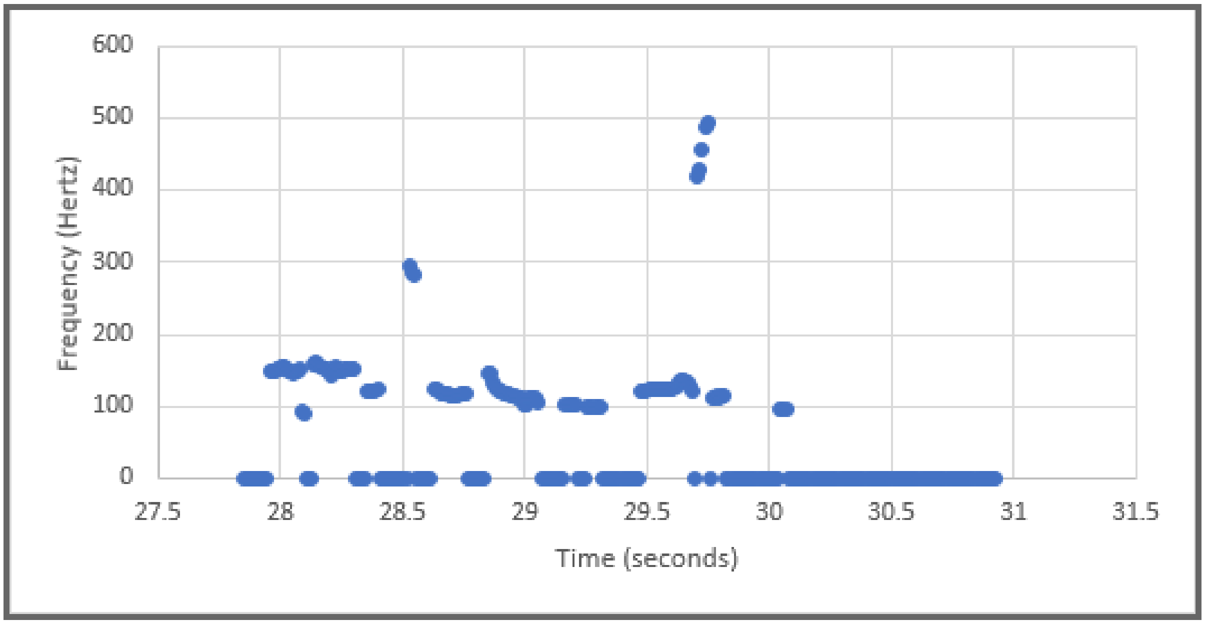


Figure 5: Pitch frequency values vs. time (seconds) of example User 2.

To begin the speech data analysis, each subject's .wav files were first loaded into Praat, a computer program software package for analysing, synthesizing and manipulating speech and other sounds (Boersma and Weenik, 2007). After the .wav file was inputted into Praat, the application required each .wav file to be analyzed within an interval of 10 seconds or less. For each of the 22 subjects, the same phrase from "The Caterpillar" passage (approximately 3.5 second duration) was chosen for the analysis of the recordings. The specific phrase being analyzed was "tick, tick tick, the caterpillar climbed slowly up the tracks".

Pitch, in speech, is defined as the relative highness or lowness of a tone as perceived by the ear, which depends on the number of vibrations per second produced by the vocal cords (Zhang, 2016). Pitch listings were analyzed as research has shown that concussed subjects are more likely to have varying pitch distributions, and impaired pitch perception and processing skills (Zhang, 2016).

Formant is defined as a concentration of acoustic energy around a particular frequency in the speech wave. There are several formants, each at a different frequency, roughly one in each 1000Hz band (Wood, 2005). Formant listings were analyzed as it has been discovered that an increase in vowel formant frequencies is consistent in persons with dysarthria resulting from traumatic brain injury. Furthermore, certain acoustic metrics such as pitch and formant frequencies can vary between non-concussed and concussed subjects (Poellabauer et al., 2015).

Figure 3 depicts an example user's speech waveform graph and spectrogram. In audio, a waveform represents the amplitude of the signal over time. In other words, it shows the maximum extent of a vibration or oscillation, basically the volume of the sound over time (Aimonetti, 2019). A spectrogram is a visual way of representing the signal strength, or "loudness", of a signal over time at various frequencies present in a particular waveform (Aimonetti, 2019). The top half of the graph showcases the raw wavelength signal of the user's speech in a 1.5 second interval, while the user is in the process of reading "The Caterpillar" passage. The bottom half of the graph is the spectrogram of the respective wavelength signal with formant frequencies overlaid in red. This image was produced by the audio processing software, Praat. This figure ultimately depicts the raw data that was gathered and synthesized into Praat, before it was processed for analysis.

The Pitch and Formant listings of every subject within the time interval of the phrase being read, was then inputted into an individual .csv file, where it was pre-processed. Figure 4 showcases a data analytics pipeline that was constructed in the Python programming language to clean the data and aid in the pre-processing and analysis. The flow chart depicted in Figure 4 was followed in order to collect, process, and analyze the speech data of the participants of the study. Each subject's speech recording was saved as a waveform audio format, loaded into Praat, and subsequently pre-processed. Each recording was similarly trimmed in order to sample an approximate 3.5 second phase from where the pitch and formant listings were obtained. Ultimately, for the purpose of retaining accurate data, the undefined values, and all time listings were disregarded from the pitch and formant listings. Once the undefined and inessential columns were removed from the listings, the data was inputted into a .csv file and imported into Python. Through the pipeline, the four statistical variables identified in the last tier of Figure 4 were calculated for the pitch, formant 1 and formant 2 values.

Figure 5 depicts an example user's data in order to provide a baseline of the pitch values. Baseline values can be categorized as values that output a "0". Unfortunately, pitch was not well defined for approximately half of the values in the pitch listings. For the purpose of aiding in the understanding of the pitch value distribution, the undefined values have been graphed as 0. From the graph, it is apparent that over 50% of the pitch values are baseless, while most other baseline values were discarded. The data points a Frequency value at about 100 Hz or greater. Similarly, for the purpose of this study pertaining to the formant listings, only Formant 1 and Formant 2 listings were used, and Formant 3, Formant 4, and time listings were disregarded. After the data was pre-processed, each individual .csv file was then imported into the python pipeline. The pipeline worked on one .csv file at a time, as it would take each .csv file and output the values of the statistical variables derived from the listings in the .csv file.

Five different statistical variables for the speech analysis were computed from the python pipeline: mean, standard deviation, variance, skewness, and kurtosis. These specific variables were chosen as they were directly applicable to speech analysis and were variables that could easily be compared to each other. The final goal of the speech analysis was to compare the statistical variable values of the non-concussed participants to concussed participants.

The mean of a set of data, is the sum of observations divided by the number of observations, while standard deviation is the measure of the amount of variation in a set of values (Petruccelli, 1999). The observations in speech are described as the pitch and formant listing values, and the average value and standard deviation of these listings was important to calculate because these values provide insight into how fair the data representation is and give a general idea of how the varying distributions of pitch and formant frequencies can be considered similar to each other, when comparing non-concussed vs. concussed subjects.

Variance is described as the measure of the divergence of a set of data from the average value (Petrucelli, 1999). Looking at this statistical variable in the pitch and formant listings was helpful in analyzing the range of pitch values and how their distribution values differ between both non-concussed and concussed subjects.

Additionally, skewness is the degree of distortion from the symmetrical bell curve or the normal distribution. It measures the lack of symmetry in data distribution (Jain, 2019). While, kurtosis is used to describe the extreme values in the tails of the distribution. It is the measure of outliers present in the distribution (Jain, 2019). Computing Skewness and Kurtosis as a statistical variable on the pitch and formant listings was important because not only does this statistical analysis give insight into how consistent and reliable the data is, and what certain parts of the data is considered to be an outlier, it also allows for discussion into whether or not the mean of the data set is a fair representation of the data.

Once all the statistical variables were computed, they were also stored into a Microsoft Excel file forming two groups: the concussed and the non-concussed subjects These two groups were compared to each other in three different ways that were all done through the help of Excel. The first way the data was analyzed was through box and whisker plots, where for each statistical variable, a comparison plot was made to aid in the analysis and understanding of the data through visual comparison. Next, to also assist in the comparison and provide a summary of the results collected and analyzed, the means and standard deviation of each of the statistical variables data was reported in a summary table where Table 3 compares and contrasts these values through statistical tests about the mean values of the distributions. Finally, in order to test if the means of the two groups are significantly different, statistical tests were conducted in the form of independent sample t-tests (Table 3). By using the Data Analysis pack on Excel, t-tests were conducted to determine if the two groups being compared had a statistical significant difference.

4.0 Analysis

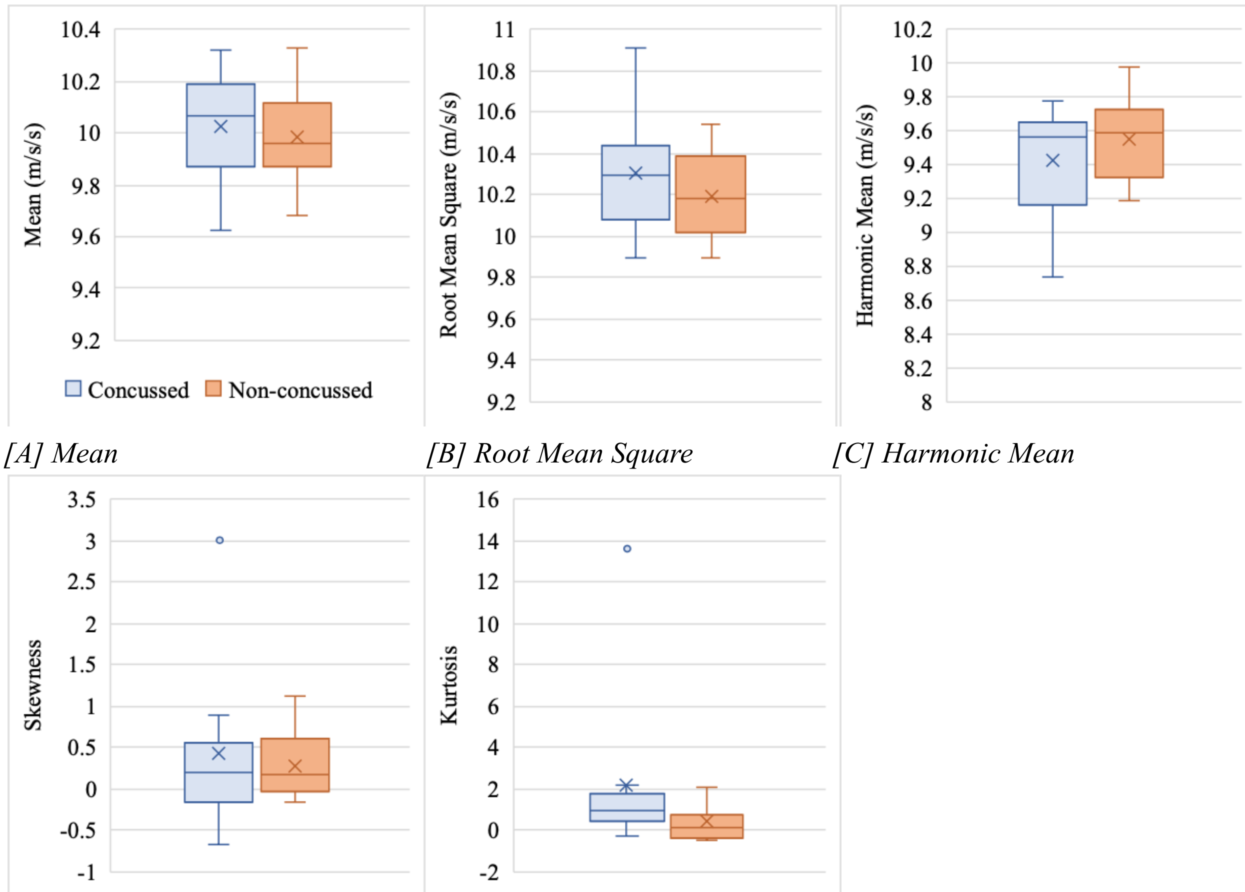
Upon collecting all of our gait and speech data from our participants, we had analyzed it by individually processing each participant's data through the respective speech and gait data pipelines. After that, we had grouped the processed participant data in Excel into the concussed and non-concussed groups to implement further analysis directed towards finding statistically significant differences between the two groups. We derived visual aids to assist the data analysis commentary. This chapter details what we discovered in the Results section, and then explains the significance behind our data analysis discoveries in the Discussion section.

4.1 Results

Gait, speech, and survey data were collected from all subjects who participated. Upon collecting the data of the participants, it was separated into two groups: concussed and non-concussed. The concussed group contained the data of participants who had indicated they had a concussion. The non-concussed group contained the data of participants who had indicated that they had not had a concussion. There were ten participants in the concussed group, and there were twelve participants in the non-concussed group. It was hoped that the groups would have the same number of members, but it was difficult to find participants with concussion histories.

4.1.1 Processed Gait Data

The processed gait data was grouped by concussed and non-concussed participants. The data analysis conducted through the box and whisker plots was to determine whether there were any significant differences between the groups in regards to the statistics drawn from them.



[A] Mean *[B] Root Mean Square* *[C] Harmonic Mean*

[D] Skewness *[E] Kurtosis*

Figure 6: Box and whiskers plots of gait for concussed and non-concussed.

As seen in Figure 6A, the mean of the mean values of both groups are close to each other as noted by the 'X' in the boxes, and the entire range of values are similar as well. There are few outliers marked to exceed the minimum and maximum lines generated by the box plot. The only graphs that have outliers are in Figure 6D and 6E. Skewness and kurtosis, which are statistical representations of the data distributions, not the data itself. The interquartile ranges, IQR, for both plots in A are different. The concussed gait plot in A has a larger range leaning towards the maximum and the non-concussed IQR is more centralized.

In Figure 6B, the averages of the RMS for both plots are different from each other with the concussed plot being higher than that of the non-concussed. This signifies in B, that there were generally higher acceleration values in the concussed group compared to the non-concussed group. The IQR in B are also different with slightly different ranges. The modes of both plots in

B were close to the means but the calculated minimum and maximum ranges of the plots in B differ with there being a much larger range for the concussed plot and a much smaller one for the non-concussed plot. This indicates there were more large RMS values in the concussed group that extended the range. The non-concussed plot in B has values barely exceeding the IQR which shows there were closer RMS values which made the IQR represent almost all of the RMS values in the non-concussed box plot.

Figure 6C shows the box and whisker plot of the harmonic means of the concussed group next to the box plot of the harmonic means of the non-concussed group. The IQR of the non-concussed plot in 6C is slightly different again between the two groups. The harmonic means of the non-concussed group were higher than the non-concussed normal means which means the typical rates/accelerations of the non-concussed group were actually higher than that of the concussed group. The mean of the non-concussed group plot in 6C was higher than that of the concussed plot and the modes were close. The calculated minimum and maximum range of the plots in 6C are different with the non-concussed plot being much smaller but set with higher values and the concussed had a large range with a minimum much further from the concussed IQR than the maximum was. This indicates the concussed plot is skewed right and that there were smaller harmonic means in the concussed group that were not as common as the larger means.

The box and whisker plot in Figure 6D shows the concussed vs non-concussed skewness values. The IQR of both plots in 6D are close to having the same range with the concussed plot being somewhat larger with a slightly smaller upper quartile and a smaller lower quartile compared to the non-concussed plot. In both plots, the values show a skewness higher than 0 which indicates the gaits of most participants were skewed right. The calculated minimum of the non-concussed plot is much smaller than that of the concussed plot in 6D, indicating that the non-concussed group was barely skewed left while the concussed participants were skewed left. There is an outlier in the concussed box plot that would mean one concussed participant's gait distribution is heavily skewed right.

In Figure 6E, both plots of the kurtosis ranges exist below 3 which signifies both plots have a lack of outliers (Jain 2019). There is one outlier in the concussed box plot at 13.57 which signifies that for one concussed participant extreme gait values were as common as normalized gait values. The IQR for the concussed group is set higher than that of the non-concussed group. The mean of the concussed plot is higher in 6E and skewed right to further signify the difference between both groups whereas the non-concussed group is skewed left.

Table 2: Gait t-Tests: two-sample assuming unequal variances

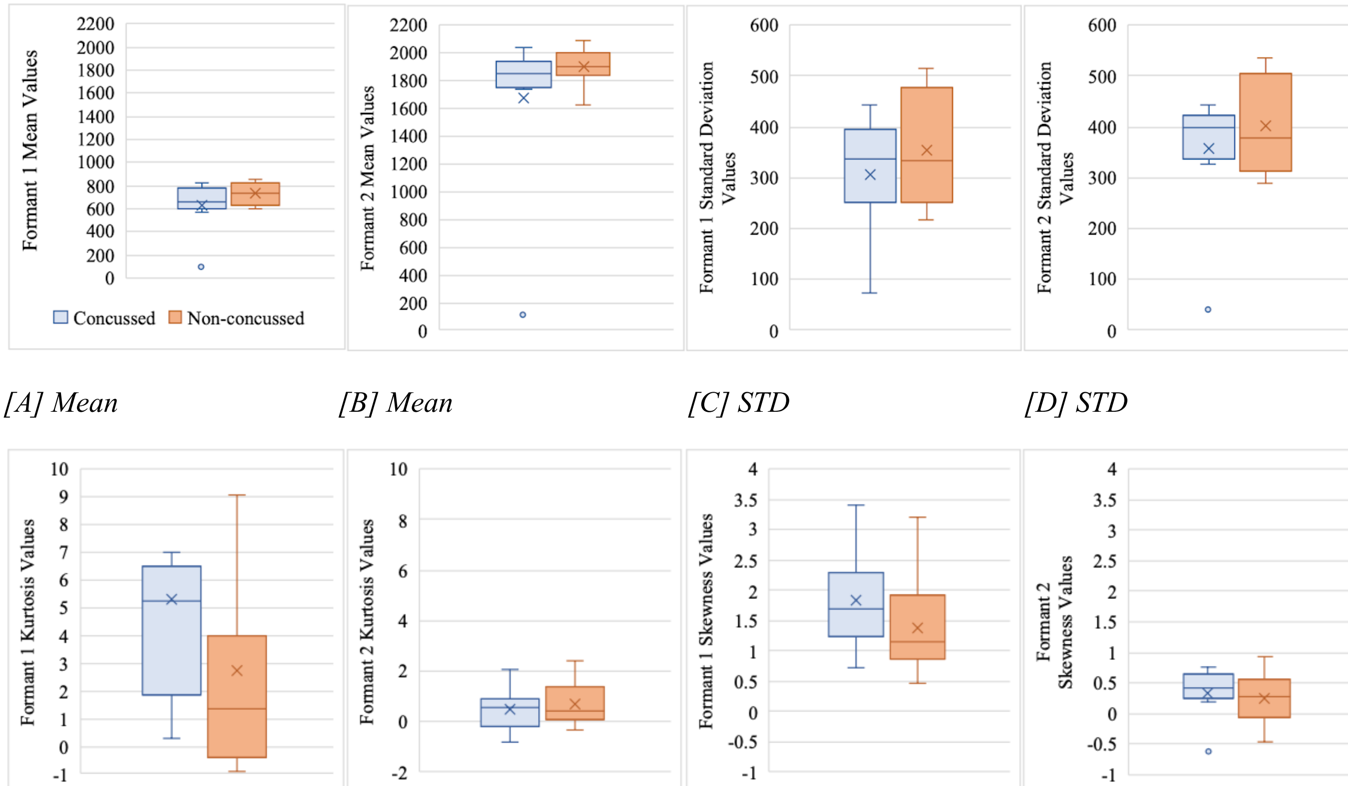
	Concussed Mean	Non-concussed Mean	Concussed std dev	Non-concussed std dev	df	t-statistic	p value
Mean	10.03	9.987	0.2297	0.1690	16	0.4612	0.6509
RMS	10.30	10.19	0.2872	0.2052	16	1.015	0.3248
Harmonic Mean	9.431	9.551	0.3396	0.2535	16	-0.9340	0.3642
Skewness	0.4188	0.2808	0.9983	0.3898	11	0.4135	0.6872
Kurtosis	2.187	0.3869	4.057	0.8286	10	1.381	0.1973

Table 2 shows the average value and standard deviations (std dev), for the statistical values compiled in the concussed and non-concussed groups. It also shows the standard deviation in those means. The two sample t-Test assuming unequal variances was done using Excel for each statistical value. Table 2 displays degrees of freedom (df), t-statistic, and the p-value for each statistical value calculated in the concussed and non-concussed groups. The Null Hypothesis is that the concussed group is not significantly different from the non-concussed group in regards to mean, RMS, harmonic mean, skewness, and kurtosis. The t-statistic was calculated in Excel where alpha was set equal 0.05, so in order for a significant difference to be proven, the resulting t-statistic has to be more than the t Critical two-tail value found at the specified df where 0.05 is the value of alpha

As seen in the last column, all of the p-values are greater than 0.05 which indicates there is no significant difference between the concussed group and the non-concussed group. Thus, the null hypothesis cannot be rejected. The skewness means were different from each other but both were ultimately positive indicating the distribution for both groups was skewed right. The kurtosis mean for the concussed group was much larger indicating a stronger distribution in the tails of the concussed acceleration values. But the standard deviation for the kurtosis was also much higher for the concussed group indicating the tail distributions were generally different from each other. The standard deviation was also higher for the concussed group in skewness as well indicating not all of the distributions were skewed right.

The t-statistics is used to determine significance between the non-concussed and the concussed group. For normal mean, the t-statistic is 0.4612 and the degrees of freedom was 16, but the t Critical two-tail was higher at 2.119 which means the null hypothesis cannot be rejected. In RMS, the t-statistic is 1.015 which is less than 2.119 at 16 degrees of freedom which also signifies the null hypothesis cannot be rejected. With the harmonic mean, the t-statistic is -0.9340 which is again less than 2.119 at 16 degrees of freedom. Next, the skewness t-statistic is 0.4135 which is less than 2.201 at 11 degrees of freedom so the null hypothesis again, cannot be rejected. And finally, for kurtosis, the t-statistic is 1.381 which again is less than 2.228 at 10 degrees of freedom which means the null hypothesis cannot be rejected across all the statistical values calculated and compared.

4.1.2 Processed Speech Data



[A] Mean *[B] Mean* *[C] STD* *[D] STD*

[E] Kurtosis *[F] Kurtosis* *[G] Skewness* *[H] Skewness*

Figure 7: Box and whisker plots for Formant 1 and Formant 2 speech metrics

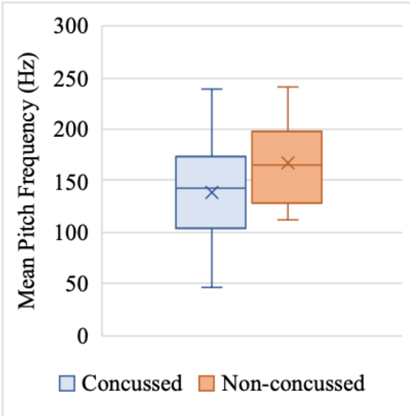
As seen in Figure 7A, the box and whisker plots of the formant 1 mean values of concussed and non-concussed subjects are similar. The most pointed difference is that the distribution for the concussed subjects is slightly lower than that of the non-concussed subjects. Furthermore, in Graph B, the formant 2 mean values of concussed and non concussed patients are similar. In graphs A and B, the concussed subjects have a greater interquartile range than that of the non-concussed subjects, and there are also outliers present in the concussed subjects data (values that are more than one and a half times the interquartile range from below q1 or above q3) indicating a higher amount of variance in the concussion subjects' values.

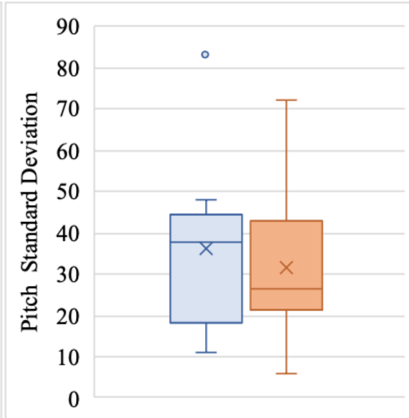
Figure 7C and 7D represent the box and whisker plots of the formant 1 and formant 2 standard deviation values of concussed and non-concussed subjects. In these distributions, a similar pattern is represented as that of Graphs A and B in that the concussed values are slightly lower than those of the non-concussed values. Moreover, the lower 25% of the mean concussed values for Graph C were markedly lower when compared to the non-concussed values. The pattern from the mean and standard deviation values, it can be summarized that the mean and standard deviation for the formant 1 and 2 values of concussed participants was lower than that of non-concussed patients.

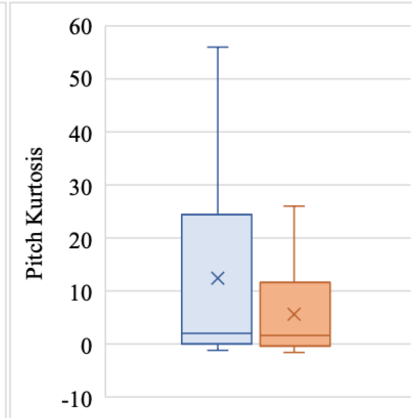
In 7E, the kurtosis values of the concussed and non-concussed values for formant 1 are characterized. Kurtosis is the combined sizes of the two tails of the gaussian, or normal, distribution curve for a series of data points (Jain, 2019). The kurtosis for a normal data series is 0, yet here, the overall kurtosis for both groups of subjects is above 0. This indicates that the datasets for both groups of subjects contained an abundance of values in the tails of the distribution curves. This signifies that there was a large spread of data present. In Graph 7F, the same data is depicted for formant 2. Here, the kurtosis values are closer to 0 indicating that the majority of collected data centered more equally around the mean value.

Finally, the skewness values of the formant 1 and formant 2 for concussed and non-concussed patients can be viewed in Figure 7G and 7H. Here, the skewness is greater than 0 in both cases, but the gap from 0 for both concussed and non-concussed is larger in formant 1 than in formant 2. Essentially, this indicates that the gaussian distribution of the data set actually has a left handed preference and a longer right hand tail, signifying that the results are right-skewed.

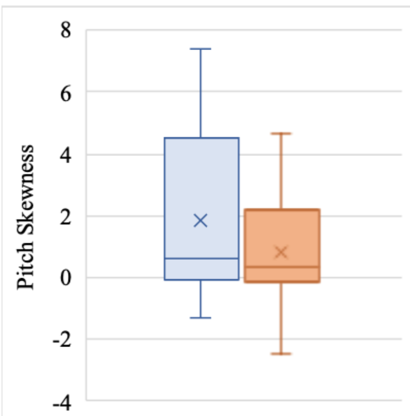
Overall, it can be summarized that the concussed values are lower than the non-concussed values for both the formant 1 and formant 2 data sets. This is supported by the lower spread of mean and standard deviation values as reflected by Graph A and B. Finally, although skewness and kurtosis values indicated a divergence from the normal distribution curve, the divergences were about similar for both formant 1 and 2 for concussed and non-concussed values, which displayed that the spread of data was similar for all of the collected results.







[A] Mean *[B] Standard Deviation* *[C] Kurtosis*



[D] Skewness

Figure 8: Box and whisker plots for pitch metrics

The pitch metric graphs display the range of the concussed subjects data to be greater than that of the range of non-concussed subjects. It can also be deduced that in the pitch variable distributions (excluding the mean value) the median is greater in concussed subjects. In Figure 8A, the mean pitch values for concussed patients was lower than that of the non-concussed patients. In 8B of the same Figure, the standard deviation values of both groups indicate that there is about an equal spread of data. Concussed patients, however, have a greater spread below the average while non concussed patients have a higher spread above the average. In 8C, kurtosis values show a spread for both datasets. Finally, in 8D both datasets have a right skew.

Overall, it can be summarized that the concussed values of the box and whisker plots for pitch metrics are lower than that of the non-concussed values. For both data groups, there was a large leftward lean in terms of the distribution of data, with most data centered toward the left of the peak distribution and a large right leaning tail. There was also a higher range in standard deviation values for the non-concussed subjects, which may have led to the mean being dragged upwards as well. It is also important to note that the concussed subjects also had more outliers.

Table 3. Formant and Pitch t-Tests: two-sample assuming unequal variances

Metric	**Concussed** *Mean (Standard Deviation)*	**Non-Concussed** *Mean (Standard Deviation)*	**df**	**t-statistic**	**p value**
Formant 1 Mean	694.7 (89.09)	728.3 (98.9)	12	-0.742	0.236
Formant 2 Mean	1867.9 (116.8)	1904.2 (123.5)	13	-0.638	0.267
Pitch Mean	151.0 (49.2)	167.0 (40.0)	11	-0.730	0.480
Formant 1 Standard Deviation	336.2 (74.0)	353.6 (103.6)	16	0.400	0.694
Formant 2 Standard Deviation	396.7 (40.1)	402.6 (88.8)	17	-0.188	0.853
Pitch Standard Deviation	37.8 (22.5)	31.8 (16.3)	9	0.582	0.574
Formant 1 Kurtosis	5.3 (5.2)	2.7 (3.8)	10	1.057	0.315
Formant 2 Kurtosis	0.4 (0.9)	0.7 (0.8)	11	-0.565	0.583
Pitch Kurtosis	12.6 (19.6)	5.9 (8.3)	7	0.791	0.454
Formant 1 Skewness	2.0 (0.7)	1.4 (0.7)	12	1.54	0.149
Formant 2 Skewness	0.3 (0.4)	0.3 (0.4)	11	0.370	0.717
Pitch Skewness	1.8 (2.8)	0.8 (2.0)	9	0.792	0.448

Table 3 summarizes the mean and standard deviation of several statistical parameters in regards to the formant 1, formant 2, and pitch data of concussed and non concussed participants. Microsoft excel t-Tests, of two-samples assuming unequal variances were conducted on each statistical value. The table displays the degrees of freedom (df), t-statistic, and p-value for each statistical value calculated in the concussed and non-concussed groups. The p-value can be utilized to indicate significance. The null hypothesis is that the concussed group is not significantly different from the non-concussed group in regards to mean, standard deviation, variance, skewness, and kurtosis

Overall, the independent t-tests display that all p values are greater than 0.05, and the t-statistic is less than the t critical two-tail value at the given degrees of freedom, indicating that the null hypothesis cannot be rejected therefore there was no statistically significant difference between the means of the concussed and non-concussed subjects.

From the table, it can be generalized that the mean and standard deviation statistical parameters of the non-concussed subjects were greater than those of the concussed subjects for formant 1, formant 2, and pitch. The variance is large for both concussed and non-concussed groups, which indicates that collected values strayed from the average. In the case of formant 2, the variance for concussed was much greater than that for non-concussed. The kurtosis values for formant 1 and pitch values are much greater than 0, which signifies a leptokurtic distribution that has a sharp peak and extended tails. The kurtosis value for formant 2 is close to 0, 0.4 for concussed and 0.7 for non-concussed, which signifies normal gaussian distribution. All skewness values are positive, which ultimately indicates right skewed distribution curves with large left hand side humps and elongated right hand side tails.

4.1.3 Survey Data

Table 4. Concussed Group Survey Responses

Question	0 = not experien-ced at all	1 = no more a problem	2 = a mild problem	3 = a moderate problem	4 = a severe problem	Total	Mean	Standard Deviation
Headaches	5	2	0	2	1	10	1.2	1.5
Feelings of dizziness	6	1	1	1	0	9	0.7	1.1
Nausea and/or vomiting	8	1	0	0	0	9	0.1	0.3
Noise sensitivity	6	1	3	0	0	10	0.7	0.9
Sleep disturbance	7	0	2	0	0	9	0.4	0.8
Fatigue, tiring more easily	5	3	2	0	0	10	0.7	0.8
Being irritable, easily angered	8	1	0	0	0	9	0.1	0.3
Depressed or tearful	7	1	1	0	0	9	0.3	0.7
Feeling frustrated or impatient	7	1	1	0	0	9	0.3	0.7
Forgetfulness, poor memory	4	2	3	1	0	10	1.1	1.0
Poor concentration	6	0	3	0	0	9	0.7	0.9
Taking longer to think	6	0	3	0	0	9	0.7	0.9
Blurred vision	8	1	0	0	0	9	0.1	0.3
Light sensitivity	6	3	0	1	0	10	0.6	0.9
Double vision	8	1	0	0	0	9	0.1	0.3
Restlessness	5	3	1	0	0	9	0.6	0.7

Table 4 displays the survey responses to the River Mead Questionnaire presented in the survey from the group of participants that had indicated they have a prior concussion history.

The participants in this group indicated they were relatively healthy with the 0 option being chosen most frequently for all questions. The lower the values, the healthier the participants were. As seen in the table, a mean of 1.0 was rarely exceeded and the few instances where choices higher than 1 were chosen can also be seen. Participants were questioned on sleep disturbance, headaches, depressed feelings, and feelings of frustration. It can be seen in the table that the answers to these questions had a stronger variety of responses. This indicates that despite a portion of the concussed group responding that they have felt some of the listed symptoms, these symptoms did not impede their normal functions since the degree of symptoms 2 was rarely exceeded.

Table 5. Non-concussed Group Survey Responses

Question	0 = not experien-ced at all	1 = no more a problem	2 = a mild problem	3 = a moderate problem	4 = a severe problem	Total	Mean	Standard Deviation
Headaches	7	2	2	0	1	12	0.8	1.2
Feelings of dizziness	8	1	3	0	0	12	0.6	0.9
Nausea and/or vomiting	10	2	0	0	0	12	0.2	0.4
Noise sensitivity	9	1	1	1	0	12	0.5	1.0
Sleep disturbance	4	6	1	1	0	12	0.9	0.9
Fatigue, tiring more easily	5	1	5	1	0	12	1.2	1.1
Being irritable, easily angered	8	3	0	1	0	12	0.5	0.9
Depressed or tearful	7	2	2	1	0	12	0.7	1.0
Feeling frustrated or impatient	7	2	2	1	0	12	0.7	1.0
Forgetfulness, poor memory	8	2	2	0	0	12	0.5	0.8
Poor concentration	8	1	2	1	0	12	0.7	1.0
Taking longer to think	9	1	2	0	0	12	0.4	0.8
Blurred vision	12	0	0	0	0	12	0	0
Light sensitivity	9	2	0	1	0	12	0.4	0.9
Double vision	12	0	0	0	0	12	0	0
Restlessness	9	1	1	1	0	12	0.5	1.0

Table 5 displays the survey responses from the group of participants that had indicated they have no concussion history. This group also indicated they were relatively healthy with the 0 option being chosen most frequently. Similar to the concussed group in Table 4, participants had a stronger variety of responses on questions regarding sleep disturbance, headaches, depressed feelings, and feelings of frustration. However, it is evident in this table that the non-concussed group had more signs of TBI symptoms since all of the mean responses are higher than that of the concussed group in Table 4. This is an interesting discovery as this distinction in group symptoms may have caused the gait and speech data between the two groups to be more similar than expected. But, the mean values were still low, never exceeding 1.

4.2 Discussion

It was hypothesized that there would be a statistically significant difference between the gait and speech results of the non-concussed and concussed groups. All the statistical values calculated and compared through the t-tests in regards to the gait study have shown that there were no significant differences between the concussed group and the non-concussed group. The survey results coincidentally compound on this information by presenting the concussed group to be showing fewer TBI symptoms than the non-concussed group.

It was predicted that for gait that the mean, RMS, and harmonic mean distributions of the concussed group would be different with a range set lower than that of the non-concussed group, which was shown to only be true with Figure 6C. But all their p-values exceeded 0.05 so the null hypothesis could not be rejected as seen in Table 2. For gait skewness and kurtosis, it was predicted the concussed group would have a lower skewness and higher kurtosis values than the non-concussed group. This would indicate the concussed group could not maintain a completely consistent gait and the non-concussed group could at a faster pace. A kurtosis value above 3 indicates concentration in the distribution curve tails to be higher than normal. In Figure 6D, the skewness distribution of the concussed group was skewed further right in comparison to the non-concussed group showing both to have higher inconsistent gaits. 6E shows both group distributions to be under 3, thus having less outliers than a normal data set. But yet again, Table 2 shows none of these gait differences were able to reject the null hypothesis.

It was hypothesized that speech would differ between the concussed and non-concussed group as well. Taking into account the results from the data that was collected in the speech study, it can be said that this hypothesis has to be rejected, due to the fact that through the thorough analysis of the data collected of the box and whisker plots, summary tables, and t-tests, the results indicate that there there was not a statistical difference between the means of the two populations. The box and whisker plots for various formant 1, formant 2 and pitch frequency metrics shown in Figure 7 and Figure 8 visually indicate that there is a noticeable distinction between the means and range distributions of both populations. Looking at both pitch and formant graphs and their respective range distributions, it is apparent that mean and standard deviation values were overall lower for the concussed subjects, versus that of the non concussed subjects. Our data does not support the hypothesis in speech. Therefore, the null hypothesis cannot be rejected as all calculated p-values are greater than 0.05.

From the survey data collected, the concussed participants do not show a high degree of symptoms of TBI, indicating that the subjects have symptoms similar to those without TBI histories and may have recovered well from their TBI. Their recoveries may have allowed for them to provide speech and gait data that does not show any sign of previous TBI.

There are many factors that could have led to discrepancy in the results of the speech study as well. All recordings of the subjects being tested were not taken on the same day, and in the same exact location, with inconsistency in volume and background noise leads to outliers and variance in pitch and formant listing frequencies. Furthermore, the varying inconsistency of the results could also be due to the fact that reading aloud and public speaking is often a difficult experience for students. For example, in select trials, subjects who have not had any prior history of concussions were struggling to read the Caterpillar passage with the same efficiency of that of the other subjects. Additionally, it must also be noted that subjects were not asked about the severity of the concussions they had suffered. The concussions recorded may have been mild and without prolonged effects that were visible in the results.

5.0 Conclusion

TBI is a major public health burden sustained by millions of people annually, and has very severe consequences if not treated quickly. In order to receive treatment, a diagnosis must first be issued by a medical professional, but the diagnostic aids commonly used by medical professionals with identifying brain injury can be slow, redundant, and intrusive, risking more damage to those injured. The ultimate goal of this study was to provide data that will aid in the development of diagnostic tools that can help diagnose concussions in a quick and non-intrusive way, and to provide preliminary analysis of the respective data as proof-of-concept for gait and speech-based indicators of concussions. Using mobile phone apps, audio/speech and gait data was collected from participants who did and did not have histories of TBI. Using original data pipelines and Microsoft Excel, the collected data was processed to derive statistical values specific to gait and speech, and then analyzed further with t-Tests, to determine if there were any statistically significant differences between the concussed and non-concussed groups.

Finding no statistically significant differences between the concussed and non-concussed groups was unexpected. However, this does not mean the project did not bear any merit. The method of data collection paired with the survey proved to be very effective in gathering an abundance of data related to symptoms of TBI in a quick and convenient fashion. Moreover, the time efficiency of the protocol was unmatched as it was initially expected that the data collection would take longer than it actually did. The .wavRecorder and Praat software for analyzing the speech data proved to be incredibly effective in providing access to processed data. The survey itself also provided an outlet to cement any gaps that may reside in the data, regarding symptoms that gait or speech cannot represent. Although the data collection fulfilled its intended goal of ease and convenience, future experimentation with the grouping and usage of the participants may be required to fully manifest the distinction between concussed and non-concussed subject groups.

Furthermore, this project became a great learning experience in regards to statistical analysis, as with no prior experience making data analysis programs, for the successful completion of this project, the ability to learn to adapt quickly was required.

There were ultimately no significant differences found in the group of participants with a history of TBI (the concussed) and the group of participants with no history of TBI (the non-concussed). This was unanimously observed over all the statistical values compiled for speech and gait. As seen in Figure 6, there were slight differences in the mean, root mean square (RMS), harmonic mean, skewness, and kurtosis through box and whisker plots comparing the gait accelerometer data of the concussed and non-concussed, but Table 2 indicates the differences were insignificant. The same observations can be made for formant and pitch frequency of the concussed and non-concussed groups' speech data in Figure 8 and 7, but there was no statistical significance found as denoted in Table 3. The results of the survey that participants completed served the purpose of explaining potential outlier data and gathering participant background data. But no irregular or outlier symptoms were found in the survey responses from both groups.

6.0 Future Work

There are several short and long term goals that will be recommended in order to assist in the future continuation of this project. The first short term goal would be to acquire IRB approval to continue testing and experimentation as soon as possible. Ordinarily, this would mandate the construction of an additional follow-up introduction, background, and procedure in order to apply for IRB approval, but only a completed procedure is necessary to commence the IRB application phase. As such, it is strongly advised to keep the introduction and background research brief.

This project has also gone through many conceptual changes over its design stages, but the one aspect that has not yet transformed is the method of collecting data from participants. That being said, if the project does experience conceptual changes, the introduction and background would require further editing and revision, which would also elongate the process of IRB approval. As such, applying for and acquiring IRB approval early on in the project schedule with the procedure alone allows for minimized stress associated with the completion of the introduction and background. This would also ensure that data collection can begin as soon as possible, once approval is granted.

Specific to experiences gathered from this project, saving time and upholding deadlines would be the largest facets of project management to improve upon. Furthermore, additional recommendations to increase the quality and time management of the project encompass the creation of a project agenda, complete with deadlines and scheduled deliverables; the conversion and enhancement of the Python pipelines constructed specifically for this project; comprehensive testing of all equipment utilized for data collection, as this guarantees their functionality and operation; and pilot studies commencing as soon as IRB approval is collected.

Based on this project's results, it is recommended that some conceptual changes be made in future efforts to address the problem. For starters, there were no significant differences found between the two groups of individuals separated by TBI. However, if the comparison could be more specific, there is a possibility that more differences could be found. Therefore, it is recommended that data collection begin in A term, from as many participants as possible.

Under this project schedule, come C term, data could be collected again from the same pool of participants. This is especially paramount for subjects that may have recovered from a concussion that they had sustained after data collection in A term, hence depicting change over time. Significance may also be found from more general and specific differences, which can be guaranteed since the comparative data originates from the same people. Moreover, given the proximity of A term to C term, the recovery time itself could be less of a factor in normalizing the data from the participant. This is another reason why acquiring IRB approval immediately is recommended: the faster data that collection can begin, the larger the pool of participants.

That being said, it is recognized that this is a difficult task to execute, so remaining organized is the key to success. It would be necessary to establish immediate contact with WPI Athletics in order to gather a pool of participants, since athletes are more likely to experience TBI through their daily sports regiments. If more athletes are recruited, the data comparison in A term and C term has a higher chance of occurring and significant differences in the statistical data collected are more likely to be present. Reusing the pipelines made in this project ensures that data analysis in C term can be managed quickly and maintaining project organization ensures that this entire system will work without putting unnecessary stress on the students and advisors.

References

Aimonetti, M. (2019, June 26). Generating Waveform Data - audio representation. Retrieved from https://matt.aimonetti.net/posts/2019-06-generating-waveform-data-audio-representation/

Bell, D. R., Guskiewicz, K. M., Clark, M. A., & Padua, D. A. (2011, May). Systematic review of the balance error scoring system. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3445164/.

Ben-David, Boaz & Icht, Michal. (2016). Oral-diadochokinetic rates for Hebrew-speaking healthy ageing population: non-word versus real-word repetition: Oral-DDK in older Hebrew speakers. *International Journal of Language & Communication Disorders*. 52. 10.1111/1460-6984.12272.

Ben-David, B. M., Moral, M. I., Namasivayam, A. K., Erel, H., & van Lieshout, P. H. (2016). Linguistic and emotional-valence characteristics of reading passages for clinical use and research. *Journal of Fluency Disorders*, *49*, 1-12.

Boersma, Paul & Weenink, David. (2007). PRAAT: Doing phonetics by computer (Version 5.3.51).

Broglio, S. P., Cantu, R. C., Gioia, G. A., Guskiewicz, K. M., Kutcher, J., Palm, M., Valovich McLeod, T. C., & National Athletic Trainer's Association (2014). National Athletic Trainers' Association Position Statement: Management of Sport Concussion. *Journal of Athletic Training*, *49*(2), 245–265. https://doi.org/10.4085/1062-6050-49.1.07

Budinger, T. (2016, April 19). An Overview of Concussion History and Needed Research. Retrieved from https://www.nae.edu/152220/An-Overview-of-Concussion-History-and-Needed-Research

Cantu, R. (2012). The Chronic Effects of Concussion on Gait. *Yearbook of Sports Medicine*, *2012*, 8–9. doi: 10.1016/j.yspm.2011.09.023

Costandi, M. (2010, November 8). Phineas Gage and the effect on personality of an iron bar through the head. Retrieved from https://www.theguardian.com/science/blog/2010/nov/05/phineas-gage-head-personality

Covassin, T., Elbin, R. J., Stiller-Ostrowski, J. L., & Kontos, A. P. (2009). Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) Practices of Sports Medicine Professionals. *Journal of Athletic Training*, *44*(6), 639–644. doi: 10.4085/1062-6050-44.6.639

Daneshvar, D. H., Nowinski, C. J., Mckee, A. C., & Cantu, R. C. (2011). The epidemiology of sport-related concussion. *Clinics in Sports Medicine*, 30(1), 1-17. doi:10.1016/j.csm.2010.08.006

Dessy, A. M., Yuk, F. J., Maniya, A. Y., Gometz, A., Rasouli, J. J., Lovell, M. R., & Choudhri, T. F. (2017). Review of Assessment Scales for Diagnosing and Monitoring Sports-related Concussion. *Cureus*. doi: 10.7759/cureus.1922

Duhaime, A. C., Beckwith, J. G., Maerlender, A. C., McAllister, T. W., Crisco, J. J., Duma, S. M., Brolinson, P. G., Rowson, S., Flashman, L. A., Chu, J. J., & Greenwald, R. M. (2012). Spectrum of acute clinical characteristics of diagnosed concussions in college athletes wearing instrumented helmets: clinical article. *Journal of Neurosurgery*, *117*(6), 1092–1099. https://doi.org/10.3171/2012.8.JNS112298

Finkelstein, E., Corso, P. S., & Miller, T. R. (2006). The incidence and economic burden of injuries in the United States. Oxford University Press, USA.

Gessel, L. M., Fields, S. K., Collins, C. L., Dick, R. W., & Comstock, R. D. (2007). Concussions among United States high school and collegiate athletes. *Journal of Athletic Training*, *42*(4), 495–503.

Graham R, Rivara FP, Ford MA, et al., editors. Sports-Related Concussions in Youth: Improving the Science, Changing the Culture. Washington (DC): National Academies Press (US); 2014 Feb 4. 3, Concussion Recognition, Diagnosis, and Acute Management. Available from: https://www.ncbi.nlm.nih.gov/books/NBK185340/

Hamilton, J. (2016, September 26). War Studies Suggest A Concussion Leaves The Brain Vulnerable To PTSD. Retrieved from https://www.npr.org/sections/health-shots/2016/09/26/495074707/war-studies-suggest-a-concussion-leaves-the-brain-vulnerable-to-ptsd

Hayes, A. (2020, November 16). Harmonic Mean. Retrieved from https://www.investopedia.com/terms/h/harmonicaverage.asp

Harrison E. A. (2014). The first concussion crisis: head injury and evidence in early American football. *American Journal of Public Health*, *104*(5), 822–833. https://doi.org/10.2105/AJPH.2013.301840

Helfer, Brian & Quatieri, Thomas & Williamson, James & Keyes, Laurel & B, Evans & WN, Greene & Vian, Trina & Lacirignola, Joe & Shenk, Trey & Talavage, Thomas & J, Palmer & Heaton, Kristin. (2014). Articulatory Dynamics and Coordination in Classifying Cognitive Change with Preclinical mTBI.. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH.

Howell, D. R., Kirkwood, M. W., Provance, A., Iverson, G. L., & Meehan, W. P., 3rd (2018). Using concurrent gait and cognitive assessments to identify impairments after concussion: a narrative review. *Concussion (London, England)*, *3*(1), CNC54. https://doi.org/10.2217/cnc-2017-0014

Jain, D. (2019, January 19). Skew and Kurtosis: 2 Important Statistics terms you need to know in Data Science. Retrieved from https://codeburst.io/2-important-statistics-terms-you-need-to-know-in-data-science-skewness-and-kurtosis-388fef94eeaa

Jones, E., & Wessely, S. (2005). War syndromes: the impact of culture on medically unexplained symptoms. *Medical History*, *49*(1), 55–78. https://doi.org/10.1017/s0025727300008280

Kenton, W. (2020, February 19). T-Test Definition. Retrieved from https://www.investopedia.com/terms/t/t-test.asp

Kleffelgaard, I., Roe, C., Sandvik, L., Hellstrom, T., Soberg, L. (2013). Measurement Properties of the High-Level Mobility Assessment Tool for Mild Traumatic Brain Injury. *Physical Therapy*, 93(7), 900–910. https://doi.org/10.2522/ptj.20120381

King, L. A., Horak, F. B., Mancini, M., Pierce, D., Priest, K. C., Chesnutt, J., Sullivan, P., & Chapman, J. C. (2014). Instrumenting the balance error scoring system for use with patients reporting persistent balance problems after mild traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, *95*(2), 353–359. https://doi.org/10.1016/j.apmr.2013.10.015

Kontos, A. P., Covassin, T., Elbin, R., & Parker, T. (2012). Depression and Neurocognitive Performance After Concussion Among Male and Female High School and Collegiate

Athletes. *Archives of Physical Medicine and Rehabilitation*, *93*(10), 1751–1756. doi: 10.1016/j.apmr.2012.03.032

Kuo, C., & Tjaden, K. (2016). Acoustic variation during passage reading for speakers with dysarthria and healthy controls. *Journal of Communication Disorders*, *62*, 30–44. https://doi.org/10.1016/j.jcomdis.2016.05.003

McCrory PR, Berkovic SF. Concussion: the history of clinical and pathophysiological concepts and misconceptions. *Neurology*. 2001;57:2283-2289.

McHenry M. (2000, March 28). Acoustic characteristics of voice after severe traumatic brain injury. *The Laryngoscope*. 110(7):1157–1161

McKee, A. C., & Robinson, M. E. (2014). Military-related traumatic brain injury and neurodegeneration. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *10*(3 Suppl), S242–S253. https://doi.org/10.1016/j.jalz.2014.04.003

Murdoch, B. E., & Theodoros, D. G. (2001). Traumatic brain injury: Associated speech, language, and swallowing disorders. Cengage Learning.

Patel, Rupal & Connaghan, Kathryn & Franco, Diana & Edsall, Erika & Forgit, Dory & Olsen, Laura & Ramage, Lianna & Tyler, Emily & Russell, Scott. (2012). "The Caterpillar": A Novel Reading Passage for Assessment of Motor Speech Disorders. American journal of speech-language pathology / American Speech-Language-Hearing Association. 22. 10.1044/1058-0360(2012/11-0134).

Peterson M, Greenwald BD. (2015). Balance problems after traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*. 96:379–380.

Petruccelli, J. D., Nandram, B., & Chen, M. (1999). *Applied statistics for engineers and scientists*. Upper Saddle River, NJ.: Prentice Hall.

Pitt, W., & Chou, L.S. (2019, May 13). Reliability and practical clinical application of an accelerometer-based dual-task gait balance control assessment. *Gait and Posture.* 71:279-283. Retrieved from https://www.sciencedirect.com/science/article/pii/S0966636219301298.

Poellabauer, C., Yadav, N., Daudet, L., Schneider, S. L., Busso, C., & Flynn, P. J. (2015).

Challenges in concussion detection using vocal acoustic biomarkers. *IEEE Access*, *3*, 1143-1160.

Rădoi, A., Poca, M. A., Gándara, D., Castro, L., Cevallos, M., Pacios, M. E., & Sahuquillo, J. (2019). The Sport Concussion Assessment Tool (SCAT2) for evaluating civilian mild traumatic brain injury. A pilot normative study. *PloS One*, *14*(2), e0212541. https://doi.org/10.1371/journal.pone.0212541

Robertson, A. (2013, October 16). Why is a widely used concussion test failing to protect athletes? Retrieved from https://www.theverge.com/2013/10/16/4842176/what-causes-impact-concussion-test-forms-unreliability.

Robbins, R. (2016) Can You Diagnose Disease from the Sound of a Voice?. Stat. Boston Globe. Retrieved from https://www.statnews.com/2016/06/29/speech-diagnosis-technology/.

Schatz, P., Kelley, T., Ott, S. D., Solomon, G. S., Elbin, R. J., Higgins, K., & Moser, R. S. (2014). Utility of repeated assessment after invalid baseline neurocognitive test performance. *Journal of Athletic Training*, *49*(5), 659–664. doi:10.4085/1062-6050-49.3.37

Schatz, P. (2010). Long-Term Test-Retest Reliability of Baseline Cognitive Assessments Using ImPACT. *The American Journal of Sports Medicine*, *38*(1), 47–53. doi: 10.1177/0363546509343805

Schatz, P., & Sandel, N. (2012). Sensitivity and Specificity of the Online Version of ImPACT in High School and Collegiate Athletes. *The American Journal of Sports Medicine*, 41(2), 321–326. doi: 10.1177/0363546512466038

Sharp, D. J., & Jenkins, P. O. (2015). Concussion is confusing us all. *Practical Neurology*, 15(3), 172–186. https://doi.org/10.1136/practneurol-2015-001087

Teel, E. F., Register-Mihalik, J. K., Blackburn, J. T., & Guskiewicz, K. M. (2013). Balance and cognitive performance during a dual-task: Preliminary implications for use in concussion assessment. *Journal of Science and Medicine in Sport*, *16*(3), 190–194. doi: 10.1016/j.jsams.2012.09.007

Turgeon, C., Champoux, F., Lepore, F., Leclerc, S., Ellemberg, D. (2011). Auditory Processing After Sport-Related Concussions. *Ear and Hearing,* 32(5):667-670. Retrieved from https://journals.lww.com/ear-hearing/Abstract/2011/09000/Auditory_Processing_After_Sport_Related.16.aspx.

Wood, S. (2005). What are formants? Retrieved from https://person2.sol.lu.se/SidneyWood/praate/whatform.html

Zhang Z. (2016). Mechanics of human voice production and control. *The Journal of the Acoustical Society of America*, *140*(4), 2614. https://doi.org/10.1121/1.4964509

Appendix

Appendix A: The Caterpillar Passage

The Caterpillar Passage: Do you like amusement parks? Well I sure do. To amuse myself I went twice last spring. My most memorable moment was riding on the Caterpillar which is a gigantic rollercoaster high above the ground. When I saw how high the Caterpillar rose into the bright blue sky, I knew it was for me. After waiting in line for thirty minutes I made it to the front where the man measured my height to see if I was tall enough. I gave the man my coins asked for change and jumped on the cart. Tick tick tick the Caterpillar climbed slowly up the tracks. It went so high I could see the parking lot. Boy was I scared. I thought to myself there's no turning back now. People were so scared they screamed as we swiftly zoomed fast fast and faster along the tracks. As quickly as it started the Caterpillar came to a stop. Unfortunately, it was time to pack the car and drive home. That night I dreamt of the wild ride on the Caterpillar. Taking a trip to the amusement park and riding on the Caterpillar was my most memorable moment ever!

Appendix B: Poster Advertising Project Research

Appendix C: Survey with the Rivermead Questionnaire

Q4 Key

Page Break

Q6 What is your sex?

- ○ Male
- ○ Female
- ○ Prefer not to say
- ○ Other (specify) ________________

Q25 What is your age?

____

Q7 What sport do you play?

- ☐ Basketball
- ☐ Baseball
- ☐ Softball
- ☐ Football
- ☐ Soccer

- ☐ Wrestling
- ☐ Track & Field
- ☐ Swimming & Diving
- ☐ Rowing
- ☐ Cross Country
- ☐ Field Hockey
- ☐ Volleyball
- ☐ SOMA
- ☐ Badminton
- ☐ Flag Football
- ☐ Squash
- ☐ Racquetball
- ☐ Wiffle Ball
- ☐ Ice Hockey

☐ Other (please specify)

____

☐ None

Q28 Have you had one or more concussions?

○ Yes

○ No

Display This Question:

If Have you had one or more concussions? = Yes

Q29 Have you had one or more concussions?

☐ Diagnosed

☐ Undiagnosed

Display This Question:

If Have you had one or more concussions? = Undiagnosed

Q21 How many UNDIAGNOSED concussions have you had?

○ 1

○ 2

○ 3

○ 4

○ > 4

Display This Question:

If Have you had one or more concussions? = Undiagnosed

Q30 When do you think your last undiagnosed concussion was?

○ less than 2 weeks ago

○ 2 weeks - 3 months ago

○ 4 - 12 months ago

○ more than 1 year ago

Display This Question:

If Have you had one or more concussions? = Diagnosed

Q8 How many DIAGNOSED concussions have you had?

○ 1

○ 2

○ 3

○ 4

○ > 4

Display This Question:

If Have you had one or more concussions? = Diagnosed

Q32 When was your last diagnosed concussion?

- ○ less than 2 weeks ago
- ○ 2 weeks - 3 months ago
- ○ 4 - 12 months ago
- ○ more than 1 year ago

Page Break

Display This Question:

Rivermead Questionnaire

0 = not experienced at all

1 = no more of a problem
2 = a mild problem
3 = a moderate problem
4 = a severe problem

Compared with how you usually feel, do you now (i.e., over the last 24 hours) suffer from:

	0	1	2	3	4
Headaches	○	○	○	○	○
Feelings of dizziness	○	○	○	○	○

Nausea and/or vomiting	○	○	○	○	○
Noise sensitivity (easily upset by loud noise)	○	○	○	○	○
Sleep disturbance	○	○	○	○	○
Fatigue, tiring more easily	○	○	○	○	○
Being irritable, easily angered	○	○	○	○	○
Feeling depressed or tearful	○	○	○	○	○
Feeling frustrated or impatient	○	○	○	○	○
Forgetfulness, poor memory	○	○	○	○	○
Poor concentration	○	○	○	○	○
Taking longer to think	○	○	○	○	○
Blurred vision	○	○	○	○	○
Light sensitivity (easily upset by bright light)	○	○	○	○	○

Double vision	○	○	○	○	○
Restlessness	○	○	○	○	○

Q7 Are you experiencing any other difficulties? Please specify, and rate as above.

	0	1	2	3	4
Other	○	○	○	○	○
Other	○	○	○	○	○

End of Block: Default Question Block

Appendix D: Participation Consent Forms

CONSENT TO PARTICIPATE IN BIOMEDICAL RESEARCH

Use of voice characteristics in concussion assessment

You are asked to participate in a research study conducted by Shikha Pandey, Mir Haaris Sultan, Adam Lammert PhD, Mark Claypool PhD, Kajal Claypool PhD from the Worcester Polytechnic Institute and the Massachusetts Institute of Technology Lincoln Laboratory (MIT LL). You have been asked to participate in this study because you are a student and/or varsity athlete at WPI participating in a sport where participants are at a relatively high risk of suffering concussion. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

- **PARTICIPATION AND WITHDRAWAL**
  Your participation in this research is completely VOLUNTARY. If you choose to participate you may subsequently withdraw from the study at any time without penalty or consequences of any kind. If you choose not to participate, that will not affect your relationship with WPI, your Varsity team, or your right to health care or other services to which you are otherwise entitled.

- **PURPOSE OF THE STUDY**
  WPI and MIT LL is currently conducting a study, to develop a mobile tool for detection and diagnosis of concussion through analysis of a brief recording of a patient's voice. This study will collect voice data for the development and validation of automated algorithms to assess the likelihood and severity of concussion and post-concussion effects where concussion is suspected.

- **PROCEDURES**
  If you volunteer to participate in this study, we would ask you to do the following things:

  You will be asked to participate in at least three voice recording and gait recording sessions: prior to, during, and after the season. In each recording session, you will be asked to perform a variety of speech tasks. Examples include reading a brief passage, repeating different syllables, producing and holding a vowel sound, describing a picture or scene, or answering questions aloud. The recording session will require no more than 3 – 5 minutes.

  Participants will walk to record motion tracking data. This could have the participants walk at their normal pace up to an estimated 100 yards in the hall outside the WPI Athletic Office. This will take 9-10 minutes. The gait will only be for walking pace and will not put the participant in any extreme conditions.

  We may ask you various questions about yourself and the nature of the event that led to the suspected concussion.

  We may ask you to perform other mild motor tasks such as rapidly tapping the phone to assess fine motor skills, walking to assess gait patterns, or performing an eye tracking task to assess oculomotor conditions.

  If you are entered into the concussion monitoring protocol because of a suspected concussion, you will be asked to participate in weekly voice recording sessions as part of the routine concussion monitoring protocol. Post-concussion recordings will be carried out at the discretion of the managing physician at WPI; you will not be asked to provide recordings if the training staff determines that you are not able to, or if doing so would pose any risk to your recovery. At the discretion of the training staff, test results from other monitor protocol tasks may be collected as part of the study.

Note that the use of voice recordings to assess concussion severity is experimental and not currently part of the concussion monitoring protocol.

- **POTENTIAL RISKS AND DISCOMFORTS**
  The procedure may involve risks that are currently unforeseeable. There is minimal risk that voice and gait recording during the concussion monitoring protocol may aggravate post-concussion symptoms. However, the training staff will have the discretion to forego or postpone voice and gait recording if (s)he determines that it will pose a risk to your recovery. You always have the right to opt out and/or decline to participate in any aspect of the study at any time for any reason.

- **ANTICIPATED BENEFITS TO SUBJECTS**
  You will have the honor of participating in a research study and will receive donuts.

- **ANTICIPATED BENEFITS TO SOCIETY**
  Your participation in this study will contribute to a national effort to improve the diagnosis and assessment of impacts of concussions on athletes, service members, and the general public. The use of voice characteristics in a mobile app to assess concussion impacts may make it possible to provide more objective, automated, and reliable concussion diagnosis in non-clinical setting than is currently available.

- **ALTERNATIVES TO PARTICIPATION**
  N/A

- **PAYMENT FOR PARTICIPATION**
  You will not be paid for your participation in this study.

- **FINANCIAL OBLIGATION**
  Neither you nor your insurance company will be billed for your participation in this research.

- **PRIVACY AND CONFIDENTIALITY**
  The only people who will know that you are a research subject are members of the research team and, if appropriate, members of the training staff involved in concussion monitoring. No information about you, or provided by you during the research will be disclosed to others without your written permission, except: if necessary to protect your rights or welfare, or if required by law. In addition, your information may be reviewed by authorized WPI representatives to ensure compliance with WPI policies and procedures.

  When the results of the research are published or discussed in conferences, no information will be included that would reveal your identity. If photographs or videos of you will be used for educational purposes, your identity will be protected or disguised. If audio recordings will be used for educational purposes, only your anonymized subject identifier will be associated with the audio.

  Your data will be encoded with an anonymous study identifier that will ensure that you cannot be identified through the use of the data. Information linking you to your identity will be stored on secure computers whose access will be limited only to authorized personnel. We intend to report findings in the medical literature and the data will be used to develop and improve automated algorithms for improved concussion diagnosis.

  Anonymized data may be used in follow-on studies and will not be destroyed.
  "Records of your participation in this study will be held confidential so far as permitted by law. However, the study investigators, the sponsor or its designee and, under certain circumstances, the Worcester Polytechnic Institute Institutional Review Board (WPI IRB) will be able to inspect and have access to

confidential data that identify you by name. Any publication or presentation of the data will not identify you."

- **CONSEQUENCES OF WITHDRAWAL**
  N/A

- **WITHDRAWAL OF PARTICIPATION BY THE INVESTIGATOR**
  The investigator may withdraw you from participating in this research if circumstances arise which warrant doing so. If you become ill during the research, you may have to drop out, even if you would like to continue. The primary investigator, Mark Claypool will make the decision and let you know if it is not possible for you to continue. The decision may be made either to protect your health and safety, or because it is part of the research plan that people who develop certain conditions may not continue to participate.

- **NEW FINDINGS**
  During the course of the study, you will be informed of any significant new findings (either good or bad), such as changes in the risks or benefits resulting from participation in the research or new alternatives to participation, that might cause you to change your mind about continuing in the study. If new information is provided to you, your consent to continue participating in this study will be re-obtained.

- **EMERGENCY CARE AND COMPENSATION FOR INJURY**
  If you feel you have suffered an injury, which may include emotional trauma, as a result of participating in this study, please contact the person in charge of the study as soon as possible.

  In the event you suffer such an injury, WPI may provide itself, or arrange for the provision of emergency transport or medical treatment, including emergency treatment and follow-up care, as needed, or reimbursement for such medical services. WPI does not provide any other form of compensation for injury. In any case, neither the offer to provide medical assistance, nor the actual provision of medical services shall be considered an admission of fault or acceptance of liability. Questions regarding this policy may be directed to WPI Health Services. Your insurance carrier may be billed for the cost of emergency transport or medical treatment, if such services are determined not to be directly related to your participation in this study.

- **IDENTIFICATION OF INVESTIGATORS**
  In the event of a research related injury or if you experience an adverse reaction, please immediately contact one of the investigators listed below. If you have any questions about the research, please feel free to contact:
  1. Shikha Pandey spandey3@wpi.edu
  2. Mir Haaris Sultan mhsultan@wpi.edu
  3. Mark Claypool claypool@wpi.edu
  4. Adam Lammert alammert@wpi.edu
- **RIGHTS OF RESEARCH SUBJECTS**
  You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you feel you have been treated unfairly, or you have questions regarding your rights as a research subject, you may contact the IRB Chair (Professor Kent Rissmiller, Tel. 508-831-5019, Email: kjr@wpi.edu) and the Human Protection Administrator (Gabriel Johnson, Tel. 508-831-4989, Email: gjohnson@wpi.edu