Relating Individual Vocal Pitch to Team Performance in A Dynamic Simulated Urban Search and Rescue Task

by

Jeska Clark

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2023 by the
Graduate Supervisory Committee:

Nancy Cooke, Chair
Robert Gutzwiller
Robert Gray

ARIZONA STATE UNIVERSITY

May 2023

ABSTRACT

Urban search and rescue (USAR) teams may use Artificial Social Intelligence (ASI) agents to aid teams in adapting to dynamic environments, minimize risk, and increase mission assurance and task performance. This thesis underlines the relationship between vocal pitch, stress, and team performance from a recent experiment conducted in a simulated USAR synthetic task environment (STE). The simulated USAR-STE is a platform to use ASI as an advisor to intervene in the human team members' cognitive processes, which aims to reduce risk to task execution and to maintain team performance. Three heterogeneous and interdependent roles interact via voice communication to search and rescue the victims: (1) medic -rescues victims and identifies the severity of injuries; (2) transporter -moves victims to their designated zone based on injury severity; (3) engineer -removes hazardous material such as rubble from a room or hallway that is blocking passage. Different speeds are associated with each role, such as medic, transporter, and engineer. Medic has a default speed; the transporter has times two over the default speed; the engineer has the slowest speed. In a total of 45 teams, three ASI conditions, manipulated based on ASI intervention communication length and frequency, were analyzed. Each team participated in two 15-min missions. The results indicate a U-shaped relationship between the transporter's pitch and a change in team performance. A possible explanation for this significance is the task and role design. The transporter may have the most central role in voice communication because when the transporter is under varying levels of workload and stress, and thus voice pitch has a complex relationship with performance for that role.

DEDICATION

First and foremost, this thesis is dedicated to my family, who instilled in me a love of learning and supported me through my academic journey. Your unwavering encouragement and belief in my abilities have been instrumental in my success.

In addition, this thesis is dedicated to those who have helped me along the way in my academic and professional careers. My deepest and sincere thanks go out to my advisor, Dr. Nancy J. Cooke, who has encouraged and generously supported me throughout my studies at Arizona State University. Without her brilliant guidance, this thesis would not have been possible. Additionally, I would like to thank, Dr. Mustafa Demir, for his encouragement and guidance throughout my research process.

I would also like to thank my thesis committee: Drs. Robert Gutzwiller and Robert Gray for their feedback and guidance.

I dedicate this work to you all.

ACKNOWLEDGEMENTS

Portions of the work presented in this thesis were supported by a research grant from The Defense Advanced Research Projects Agency (DARPA) - Artificial Social Intelligence for Successful Teams (ASIST) program award HR001119C0130 (Program Manager: Dr. Joshua Elliott). The ASU research team who conducted the Study 3 experiment: Scientist: Dr. Lixiao Huang; Graduate Students: Christopher Corral, Matthew Willett, Myke Cohen, Xiaoyun Yin, Aaron Teo, Verica Buchanan; Undergraduate Students: Arya Chethikattil, Noel Campbell, Jessica Lee.

TABLE OF CONTENTS

Page

CHAPTER

CHAPTER Page

APPENDIX

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In urban search and rescue (USAR) tasks, USAR teams are required to navigate dangerous environments, understand basic medical procedures to aid victims, and clear any foreign hazards that hinder the success of a time-critical mission (Murphy et al., 2008). Occasionally, these attributes are called to perform tasks with limited resources and assistance. Utilizing machines like an Artificial Social Intelligence agent (ASI; Williams et al., 2022) can help USAR teams adapt to dynamic environments, minimize risk, and increase mission assurance task performance. In a complex task environment, the margin for error during a mission is minimal, and team cohesion is vital for optimal performance. Tasks within a USAR mission include transporting subjects, attending to medical requirements, and effective navigation to the designated safe zone. A time-critical team task requires faster cognitive processes among the team members, including communication (Demir et al., 2017; McNeese et al., 2018), coordination (Demir et al., 2019; Gorman et al., 2019), and decision-making (Forstmann et al., 2008). The same processes are required for team effectiveness in human-AI-robot teams (HARTs).

Team Effectiveness

Team effectiveness refers to a team's capacity to achieve its goals and objectives, which leads to improved outcomes for individuals and teams (Kozlowski & Ilgen, 2006). A theoretical framework for team effectiveness is built on the input–process–output (I-P-O) heuristic (Salas et al., 1992), and refers to team organization during a task in which team members combine individual or shared resources to complete task demands (Kozlowski & Ilgen, 2006). An effective team requires that cognitive, motivational, and behavioral resources are aligned with a task objective (National Research Council, 2015). Team effectiveness, therefore, varies with individual knowledge, communication, coordination, skills, abilities, and other characteristics within the team, collectively referred to as a psychological state. The psychological state, particularly the emotional state, can be captured via vocal pitch as an indicator of emotional stress (Pisanski et al., 2016; Scherer, 1986; Van Puyvelde et al., 2018; Zhang, 2016). The vocal pitch and emotion change during the voice production process (Bunn & Mead, 1971a), in which the average pitch has been correlated with emotional changes (Banse & Scherer, 1996; Fairbanks & Pronovost, 1939).

This thesis addresses the relationship between vocal pitch, stress, and team effectiveness (i.e., team performance) from a recent experiment conducted in a simulated USAR synthetic task environment (STE) (Freeman et al., 2022). The simulated USAR-STE is a platform to use ASI as an advisor to intervene in the human team members' cognitive processes, which aims to reduce risk to task execution and to maintain team effectiveness.

Artificial Social Intelligence

Social intelligence is the capability to infer and predict human intentions, emotions, and actions (Frith & Frith, 2006). ASI utilization is increasing in most areas of industry and defense; however, enhancing the robustness of ASI social intelligence is required to effectively support dynamic human needs and to adapt to dynamic task environments. An adaptable and socialized ASI can intervene at the right time (Corral et al., 2021) and support smooth teamwork (Murphy et al., 2008). Designing effective ASIs to understand human cognitive behaviors (e.g., communication and coordination, decision-making) to be more adaptive to the team task environment and maintaining team effectiveness (i.e., team situation awareness and performance) constitutes a central component to ongoing ASI research.

The classical models for explanations for human behaviors (Von Neumann & Morgenstern, 2007) have a theoretical approach to humans and decisions. An alternative approach to transitioning an ASI to analyze human behaviors with a plan uses a Human Mental Model for the ASI agent (Chakraborti et al., 2017). According to Chakraborti (2017), the Human Mental Model is defined as an [AI] agent acknowledging the needs given the human's psychological state. The "Sense-Model-Act- Plan" outlines an [AI] agent with the ability to sense team dynamics, plan based on team behaviors, and contribute to team actions and goals to predict team social behavior. The model can capture the mental and physical state of an effective team. These models are a guideline for creating a socialized intelligent agent; however, ASI does not consider previous human exposure to real-world experiences (Chakraborti et al., 2017). As an extension of this concept, this thesis investigates the question of whether by utilizing psychophysiological metrics (i.e., vocal pitch) for ASI development will be able to better understand the needs of team members in team-based tasks.

The Goal of the Study

The present experiment (referred to as ASIST Study 3) is part of a larger research effort with DARPA ASIST, to develop an ASI agent to improve team effectiveness in dynamic task environments (Freeman et al., 2022). Observations from this study and participant feedback have pointed to a significant challenge for ASI in its ability to consider team environmental and social contexts (Chakraborti et al., 2017). Some challenges included understanding when a participant needed assistance, understanding the contextual meaning of phrases, and providing interventions at the right time.

This thesis examines teammates' vocal pitch during team tasks as an indication of stress, and in turn, team effectiveness (i.e., performance) in the context of Study 3. These results may support ASI's use of vocal pitch data during the task missions to intervene at an appropriate time and place to support team effectiveness in stressful environments. Using vocal pitch may be useful to understand the individuals' physiological state or if a speaker is experiencing emotional stress based on the vocal frequencies (Cabrera & Baez, 2011; Sondhi et al., 2015).

Rationale

The relationship between voice pitch and team effectiveness under emotional stress is presented from a variety of perspectives in literature. For instance, an increment in workload during a task can increase emotional stress on humans, as well as impact psychological and physiological states (Warm et al., 2008). The concept of stress, referred to as *emotional stress,* is defined by the Oxford English Dictionary (*Stress, n.*, n.d.) as "mental or emotional strain or tension." In this regard, emotional stress is a state in which the physiological system is disorganized, which results in decreased well-being (Gaillard, 1993). The existing work shows that voice pitch increases during stressful scenarios (Ruiz et al., 1996). In this previous study, they found that an emergency appears to trigger a spontaneous dramatic increase in pitch in the voice. A laboratory stress situation and a real-life emergency were compared in this study. The voice responses observed by the pilot and co-pilot were interpreted as a potential result of the professional function and expected agency and coping during the situations. In addition, a collaborative review article (Van Puyvelde et al., 2018), examines the speech parameters encountered that affect the voice output of the different types of load (cognitive, emotional, physical, and environmental). One of the results shows that vocal fold reactivity during phonation (i.e., speaking) is affected by emotional and cognitive load (2018). According to the article, this may be because cognitive control mechanisms compete for resources with emotional and sensory mechanisms (i.e., anxiety, fear, and excitement); cognitive demands reduce the influence of high-arousal anxiety.

Summary

This thesis is intended to examine the feasibility of using human vocal pitch as part of a larger set of psychophysiological metrics (i.e., heart rate, EEG, skin sweat response) for application to a future implementation of ASI that can understand human emotional states.

This study empirically contributes to the relationship between vocal pitch, stress, and team performance in a simulated USAR task environment. Currently, there is not any prior work examining stress vocal pitch with simulated search and rescue teams. However, a review of vocal pitch literature and a description of the experimental task of ASIST Study 3 within the synthetic USAR environment follows to provide more insights on the topic.

CHAPTER 2

BACKGROUND

The state of the literature on stress, cognitive workload, and voice pitch has been reviewed to understand the relevance of vocal patterns with stress, how stress affects teams, and the application of physiological measures in USAR environments. Studies were identified via literature searches in the Google Scholar, ScienceDirect, and the Arizona State University (ASU) research databases using the keywords "voice" in combination but not limited to "stress," "voice analysis," "team," "pitch," "workload," "emotion," "cognitive," "team effectiveness", "human performance." Currently, there is a lack of literature pertaining to the direct juxtaposition of USAR and voice pitch; however, there are present studies that discuss complex team environments.

Cognitive Workload

In complex teamwork environments, the cognitive workload can be dynamic and potentially dangerous, resulting in performance decrements (Biondi et al., 2021). This process is particularly relevant for search and rescue (USAR), aviation, and law enforcement environments that pose a life-threatening outcome. When workloads exceed the demands of individual tasks, performance and team effectiveness may suffer (Hart & Hauser, 1987).

Furthermore, requirements in team coordination increase individual workload beyond what is inherent in individual task demands (Urban et al., 1995). For instance, studies on cognitive workload and communication (Salas et al., 2008; Khawaja et al., 2012; Roßnagel, 2000) suggest that communication tends to shift from baseline communication when increasing the intensity of workload at various time intervals. In addition, team decision-making (Urban et al., 1996; Zavala et al., 2018) under a high cognitive workload can be impaired if the task load is outside the limits of the team. In Palada et al. (2018) measured accuracy target detection task that presented a UAV camera ocean view with multiple ships synchronously passing the camera's field of view. The participant had to classify the ships within the masked ship. The results show that increased speed associated with decreased decision time also results in decreased accuracy, and that the quality of evidence is related to the time spent encoding a stimulus. The results from this study show that high workload can impair task objectives.

In scenarios requiring a high cognitive workload, it is vital to control emotions and stress to maximize performance. The Yerkes-Dodson model (Teigen, 1994) suggests an inverted U-shaped relationship between performance and arousal. According to the model, increasing arousal can help optimize performance, but only to a certain point at which further increases in arousal will degrade performance. The team's performance and proficiency at handling a high cognitive workload within a certain time frame depends on the team's structure and experience within the task parameters. Stress results from levels of arousal beyond optimal and is often associated with high cognitive load, which can be examined through psychophysiology (Jerčić et al., 2020).

Voice Pitch

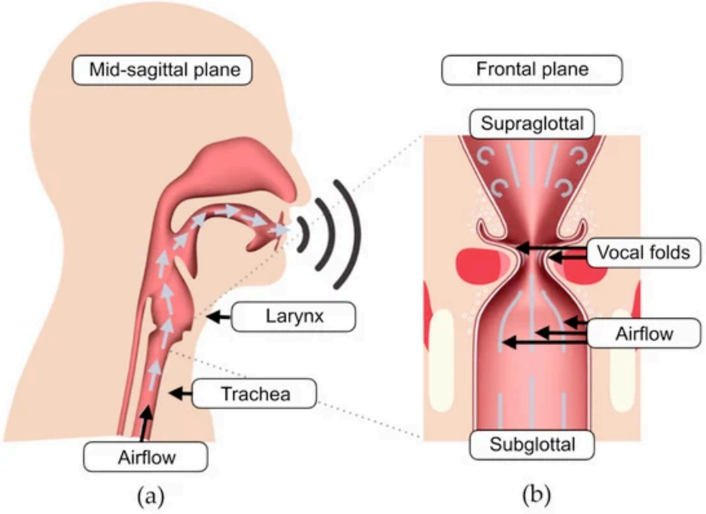
The field of psychophysiology focuses on the relationship between psychological influence and resulting physiological responses (Andreassi, 2013). The importance of psychophysiological measures (i.e., heart rate, skin conduction, and blood pressure) is in providing unique information about an individual to aid in a dynamic environment (2013). For instance, evaluating a pilot's response to a multitasking scenario using psychophysiological measures may highlight, e.g., insufficient human factors considerations in the system design.

Campbell and Ehlert (2012) and Crosswell and Lockwood (2020) collected qualitative questionnaire data, as well as physiological responses to measure individual stress levels. Participants were asked survey questions after the task was administered in order to determine whether they were experiencing stress consistent with their heart rate variability. On the survey, participants indicated little to no stress, although their heart rates during the task indicated higher fluctuations than their baselines. These studies show that relying solely on qualitative data has limitations, as an individual may be prone to over- or underestimation. Another advantage of physiological measures is that they are taken continuously without task interruption. An in-depth understanding of the psychophysiological effects of stressful events may provide insight into emotional and psychological states.

In relation to human performance, the voice output is a psychophysiological response that is part of the human integrative psychophysiological stress system, and stress reactivity is the complex integration of the sympathetic and parasympathetic nervous systems (Thayer & Lane, 2009). Previous work has shown that voice metrics can indicate if a speaker is experiencing emotional stress based on the vocal frequencies measured using a vocal stress analyzer (VSA; Cosetl & Lopez, 2011; Sondhi et al., 2015). A VSA provides metrics of the involuntary psychological, stress-related response from a person's voice (Cosetl & Lopez, 2011). In this study, VSA measured the fundamental frequency in a series of interviews from participants to detect lies or feeling ashamed. The results show that VSA technology can identify stress in the voice better than polygraphs.

Human voice with stress. To allow ASI to potentially intervene using voice pitch as a signal of stress it is important to understand how the human voice operates under stress. Emotional stress impacts voice production by affecting three processes: breathing, phonation, and resonance (Kreiman & Sidtis, 2011). Figure 1 illustrates the phonation process, where the vocal folds produce sound. Subglottal pressure determines vocal loudness as part of this process, while the cricothyroid muscle's adduction impacts the vocal-fundamental frequency, f_0, i.e., voice pitch (Van Puyvelde et al., 2018; Zhang, 2016). Resonance relates to vocal vibrations, which filter the frequencies for natural speech (Story et al., 2018). The vocal tract comprises multiple containers of air that vibrate at specific pitches, and their resonant frequencies change by altering the shape and formation of the mouth, throat, and lips (Gage & Baars, 2018). When a balance is achieved between a small glottal opening and low airflow, phonation can be preserved (Bunn & Mead, 1971).

Figure 1. A simplified mid-sagittal and frontal plane depiction of air form voice production (Thornton et al., 2019).



During a high cognitive workload, the resonance process in speech is influenced. A study found that male participants with both low and high anxiety traits showed higher f_0 under high levels of cognitive workload (Tolkmitt & Scherer, 1986). The average pitch range for male voices is 60–180 Hz, and the average pitch range for female voices is 160–300 Hz (Re et al., 2012). There are two types of voice change that are directly indicative of psychological stress: gross change (voice tremor) and voice quality (Cosetl & Lopez, 2011). A third signal category exists in the human voice, which is related to the second type of voice change. This signal category is an infrasonic signal, which is one of the more significant voice indicators of psychological stress (2011). These psychoacoustic features, such as pitch changes, voice quality, energy level, and articulation, are related to emotional state (Sobol-Shikler, 2009) due to the differentiation of emotions at the level of the autonomic nervous system. Voice can therefore indicate emotional stress-based frequencies that are associated with performance.

Research Questions and Hypotheses

This study consists of the following research questions (RQs).

RQ1: *how does vocal pitch differ across routine and novel (i.e., pre- and post-perturbation) task periods for individual roles in high-and-low-performing teams?*

H_1**:** *There will be an increase in voice pitch between pre-and post-perturbation time intervals.*

Rationale**:** The literature associates a sudden change (such as a perturbation) with a psychophysiological response of high voice pitch (Pisanski & Sorokowski, 2021). In a change from routine operations, individuals may experience stress because perturbation is a signal blackout resulting in failure to see teammates and victims on the map. There may be a difference between the voice *pre-to* and during (*post*) the perturbation in Mission 2.

RQ2: *What portion of the team's performance is explained by each role's pitch change after a perturbation?*

H_2: *There will be an increase in vocal pitch after a perturbation is associated with a decrease in performance; the pitch change will be different for each role.*

Rationale**:** Based on the theoretical framework for team effectiveness I-P-O heuristic (Salas et al., 1992), team dynamics may change to meet the goals and objectives of the team and individual roles, which may be reflected in the voice pitch. There is a difference in speed, tasks, and individual information across roles in the experiment. The voice pitch change may be an indication of cognitive load in conjunction with emotional stress (Brenner et al., 1994). Pisanski and Sorokoski (2021) found a significant increase in f_0 and intensity when a difficult task level was compared to an easy task level. Therefore, in this study, it is assumed that identifying stress during the transition from low to high workload situations can enhance ASI's ability to understand the psychological state of human subjects and ultimately intervene. In this thesis, it is hypothesized that the vocal pitch of team members will be associated with team performance in team roles during simulated USAR tasks.

CHAPTER 3

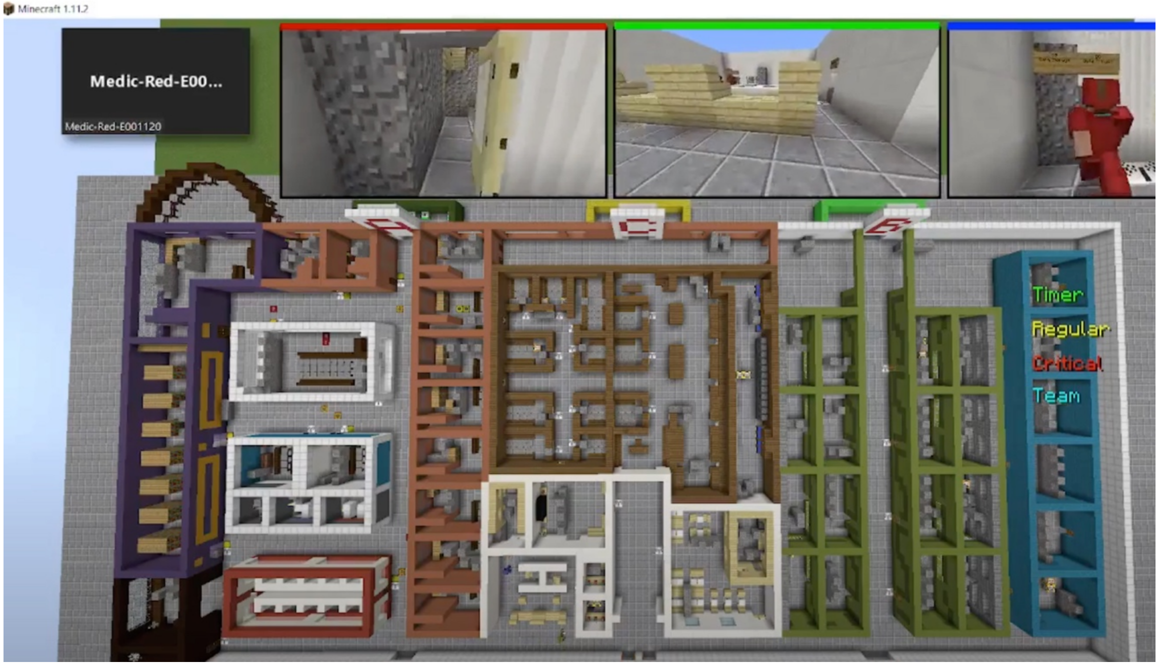
METHODS

This study is part of a large research program sponsored by DARPA (HR001119C0130). The main goal of the ASIST (Artificial Social Intelligence for Successful Teams) program is to develop ASI that intervenes in team coordination behavior by advising the team members on teamwork needed to adapt to the dynamic task environment (Freeman et al., 2022). To achieve this goal, ASIST evaluates ASI agents' interventions to all-human team interaction (i.e., communication and coordination) to improve team effectiveness.

Simulated Minecraft Synthetic Task Environment

This study was conducted remotely using an ASU Minecraft Server. The experiment was controlled by an Experimenter-Narrator, who instructed the participants through the study, and the Experimenter-Administrator, who oversaw the Minecraft training, ran the testbed and addressed any technical difficulties. The team task map was designed to resemble a corporate office building in a 50 x 50 radium in Minecraft. Figure 2 illustrates the divisions of the map, which was divided into 63 rooms, including one large library, three conference rooms, a large cafeteria, two computer rooms, and 26 storage areas, see Figure 2. There were two types of victims, regular and critical victims, with subcategories of A (bleeding), B (broken bones), and C (critical) damage types. Regular victims were worth 10 points, and critical victims were worth 50 points.

Figure 2. An overhead screenshot of the admin view during Minecraft (Freeman et al., 2022).



Team Task Roles

The Minecraft task environment consists of three heterogeneous and interdependent roles that interact via voice communication: (1) *medic,* who rescues victims and identifies the severity of injuries; (2) *transporter,* who transports victims to their designated zone based on injury severity; (3) *engineer* who removes hazardous material such as rubble from a room or hallway that is blocking passage. All team members could transport victims, but not as quickly as the transporter. There are different speeds associated with each role, such as medic, transporter, and engineer. *Medic* has a default speed; *transporter* has times 2 over the default speed; and *engineer* has the slowest speed.

Design

The experimental design for Study 3 contained 8 advisor conditions (between-subjects) and 2 missions (within-subjects). The conditions included six types of ASI advisors which differed based on developer organization, teams with no advisor, and teams with a human advisor. Each team was randomly assigned one of those conditions to aid in the two official missions. In this thesis, only Mission 2 was considered due to a significant perturbation, and teams (high or low performing) were selected from three ASI conditions. These three ASI conditions include three ASI agents that were selected to move forward in the program:

1. Carnegie Mellon University Robotics Institute (CMU) ASI: sent the messages in terms of moderate frequency and moderate message length; CMU-ASI intervenes to correct team processes, and interventions are focused on the environment, deadlines, task location, and encouragement to communicate (Freeman et al., 2023). Among the high- and low-performing teams, the CMU agent was assigned to 2 teams.
2. Dynamic Object Language Labs (DOLL) ASI: sent messages of low frequency and long message length; DOLL-ASI focuses on interventions that build relationships, guide tasks early, and improve team processes at an individual level (2023). Among the high- and low-performing teams, the DOLL agent was assigned to 13 teams.
3. Charles River Analytics (CRA) ASI: sent the messages with high frequency and short message length. CRA-ASI distributes tasks across a team by helping the team align individual skills with tasks. Interventions are focused on strategy, resource use, and motivation (Freeman et al., 2023). Among the high- and low-performing teams, the CRA agent was assigned to 11 teams.

In the experiment, there were two missions (within subjects). Mission 1 was not considered because its perturbation did not interfere with locating victims, while Mission 2's perturbation interfered, but both missions had the same number of victims (each 50: total 100 victims). In this research, Mission 2 voice pitch and team performance were examined across routine and novel time intervals (i.e., pre-post perturbation). Routine and novel time intervals are defined as the time before and after the perturbation (but including the planning process, which is the first two minutes at the beginning of the mission); routine time interval is voice pitch before the perturbation from the first five minutes of the mission, and the novel time interval is voice pitch after the perturbation occurs from minute twelve until the end of the mission.

Participants

A total of 360 participants from ASU and across the U.S. were split into 120 teams of three participants and completed Study 3. Participation required normal or corrected-to-normal vision, fluency in English, and proficiency in the video game Minecraft. Participants ranged from 18 to 66 years old (M_{age}= 23.5, SD_{age}= 6.40). However, the present work considers only 26 teams from the three ASI conditions (CMU, DOLL, and CRA). Participants were randomly assigned to the roles of Medic, Transporter, and Engineer. Each participant was compensated US $35.00 for their participation at the end of the debriefing.

Procedure

Each team participated in two competency tests (skill at utilizing basic controls and movement in Minecraft), a 10-min hands-on training mission, and two 15-min official missions. Within the missions, in Mission 1, there was no perturbation. Mission 2 perturbation included a blackout signal which hinders navigation (i.e., Global Positioning System - GPS signal) visibility within the Minecraft environment, so team members cannot see where their teammates and marked victims are on a map. Participants communicated through voice on ZOOM software (*Zoom Chat*, n.d.) to their teammates and experimenters. During the missions, teammates used individual call signs to communicate. The goal of the team task is to search for victims, stabilize them, and move them to the correct zone in the task map during each mission.

Surveys were taken when participants started the study, and after each mission, the participants completed a reflection survey about the mission outcome (i.e., strategy, team coordination, leadership), and a debriefing was given at the end of the study.

Measures

This study collected the following measures during the task: trust in an ASI, team performance, team process measures (such as team coordination and communication behaviors), information sharing, individual stress, cognitive workload, and demographics. The following measures were collected to address RQ1-2 of this thesis.

Vocal Pitch: Each participant communicated vocally with their teammates. For each mission, communication was recorded via ZOOM (*Zoom*, n.d.) and Google Meets (*Google Meet*, n.d.), which provided global and isolated audio of each team member's voice. The fundamental frequency (f_0; used to identify vocal pitch) was extracted from the audio data in the two missions using OPENSMILE software (*OpenSMILE 3.0 - AudEERING*, n.d.). The vocal analysis was carried out only on the audio data extracted from Mission 2. Matlab was subsequently applied to audio recordings to analyze the pitch deviations throughout Mission 2 (pre-and-post perturbations). The pre-perturbation interval ran from 2 minutes to 7 minutes into the mission, and the post-perturbation interval ran from 10 minutes to 15 minutes into Mission 2. Thus, for RQ1 and RQ2, the vocal pitch was divided into pre-and-post-perturbation time intervals. The vocal pitch was collected every millisecond, was summed within each perturbation time interval, and then divided by the total pre- or post-duration milliseconds to get an index of mean pitch per millisecond for that interval.

Team Performance: was calculated at the end of each mission as a cumulative score ranging from 0-900 points. The score was calculated according to the team's efficiency in maneuvering and placing the victims in the correct zone within the 15-minute time frame in each mission. Regular victims were worth 10 points, and critical victims were worth 50 points.

High-and-low-performing teams: The high-and-low-performing teams used for RQ1 were chosen based on the criteria of 2-Standard Deviations (SD) above and below the mean Mission 2 team performance score of 45 teams in the experiment using three condition ASI (DOLL, CRA, CMU). A total of 26 teams resulted, with 11 high-performing and 15 low-performing teams. High and low-performing teams were binary-coded as 1 and 0, respectively.

Pre- and post-perturbation performance: In this study for RQ2, the team performance score was divided into two separate scores for pre-and post-perturbation. During the experiment, the pre-perturbation interval ran from 2 minutes to 7 minutes into the mission, and the post-perturbation interval ran from 10 minutes to 15 minutes into Mission 2. The pre-perturbation score was the score achieved for that team by the end of the pre-perturbation period, and the post-perturbation score was calculated by subtracting the score at the 10-minute mark from the Mission 2 score.

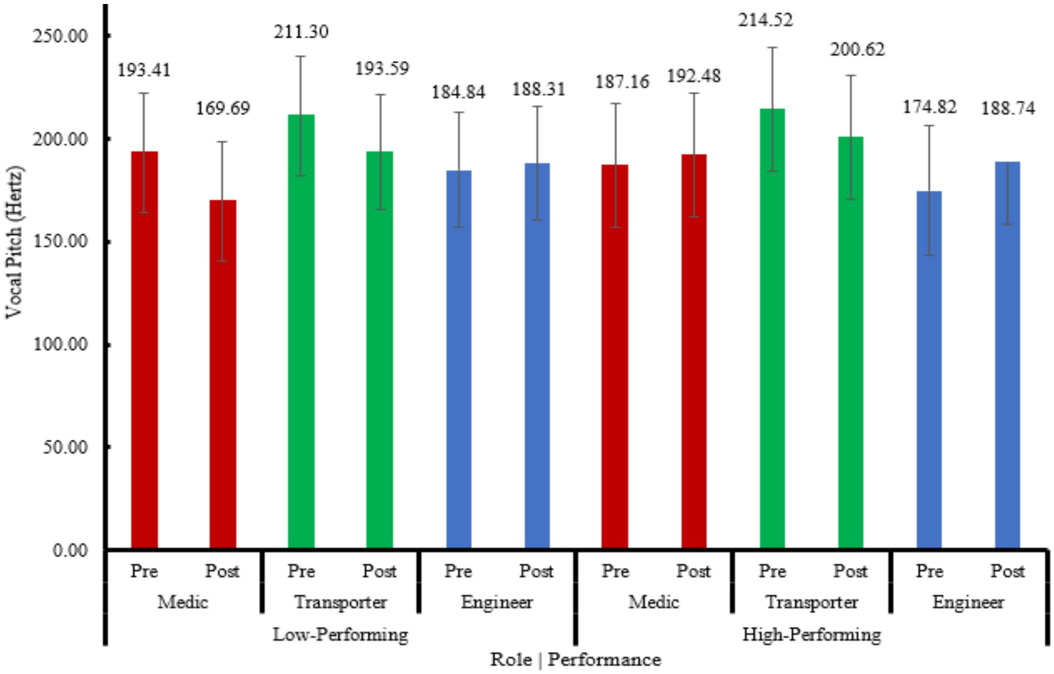
CHAPTER 4

RESULTS

M-Level Mixed Analysis of Variance (ANOVA)

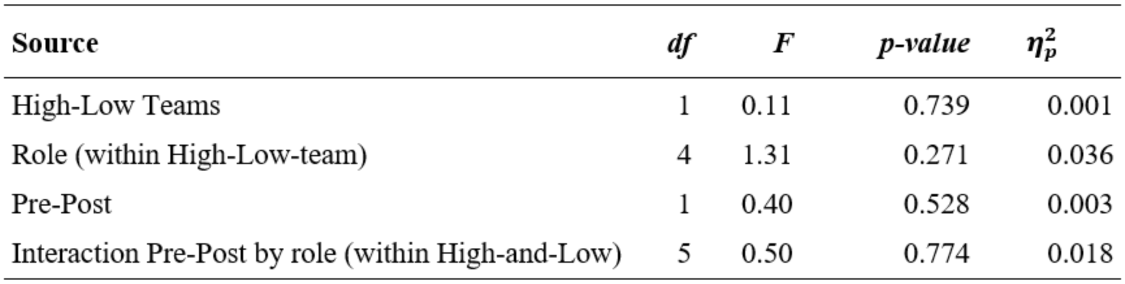
In the context of **RQ1:** *how does vocal pitch differ across routine and novel (i.e., pre- and post-perturbation) task periods for individual roles in high-and-low-performing teams,* three-level nested mixed Analysis of Variance (ANOVA; *14.2 - The General M-Stage Nested Design | STAT 503*, n.d.) was conducted. The descriptive statistics state the pitch mean for the medic role during pre-perturbation (*M*=183.08, *SD*= 60.20) and in post-perturbation (*M*=173.68, *SD*= 63.86). The pitch mean for transporter role pre-perturbation (*M*=204.66, *SD*=74.28) post-perturbation (*M*=196.83, *SD*=57.62). The average pitch for the engineer role for pre-perturbation (*M*=173.49, *SD*=58.91) post-perturbation (*M*=188.51, *SD*= 37.91).

Figure 3. The pre-and-post perturbation vocal pitch for each role.



The nominal high-low-team nominal performance score was a predictor, along with three roles nested within high-low teams. Time (pre-and-post-perturbation) was a within-subjects variable. Voice pitch was the dependent measure. The ANOVA was conducted via Statistical Package for Social Sciences Version 26 (*SPSS Statistics* | *IBM*, n.d.). The findings show that the interaction effect of pre-and-post perturbation by high-and-low performing team was not significant, *F*(5,140) = 0.503, *p* = 0.774, p2=0.184, showing that *no difference exists in vocal pitch in pre- and post-perturbation during Mission 2*; rejects **H_1.** The results depicted in Table 1 illustrate that the role within (nominal team performance) effect was not statistically significant, *F* (4,140) = 1.30, *p* = 0.271; Table 1.

Table 1. Test of between-and-within-subjects effects.

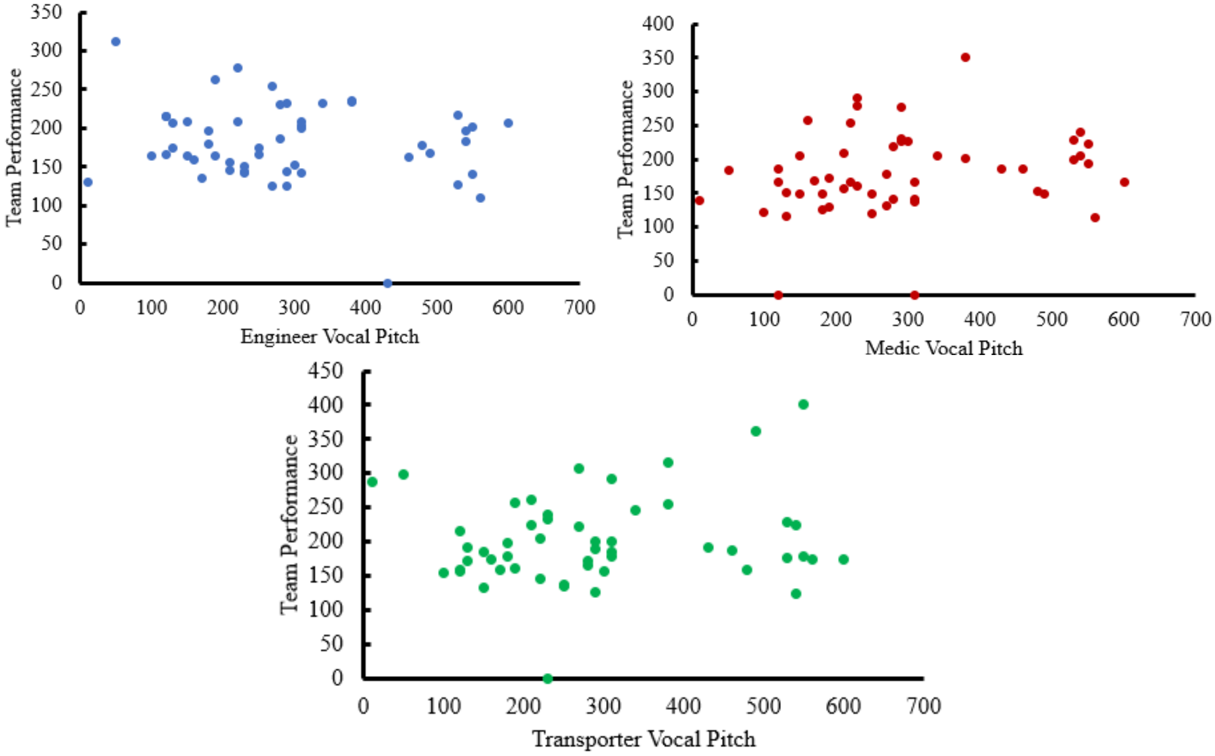


Source	***df***	***F***	***p-value***	η_p^2
High-Low Teams	1	0.11	0.739	0.001
Role (within High-Low-team)	4	1.31	0.271	0.036
Pre-Post	1	0.40	0.528	0.003
Interaction Pre-Post by role (within High-and-Low)	5	0.50	0.774	0.018

Linear Mixed Modeling

A linear mixed effect modeling procedure (Mirman, 2014) was conducted to address **RQ2:** *What portion of the team's performance is explained by each role's pitch change after a perturbation;* The following scatter plots illustrate the relationships between each role's pitch and team performance.

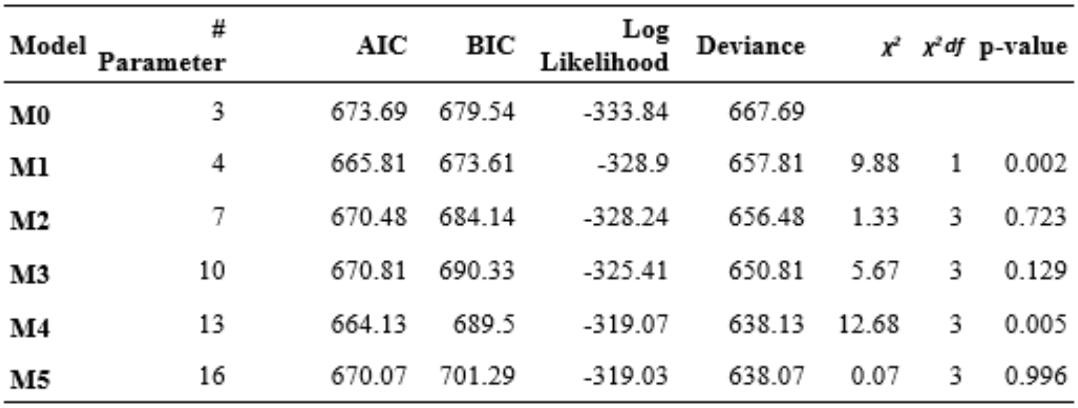
Figure 4. Scatter plots illustrate the relationship between team performance scores (pre-and-post) with 26 teams and task-role vocal pitch.



The team performance score was modeled by a linear mixed model with fixed effects of each role's pitch and their quadratic effect variables. For the final model, statistical significance (*p*-values) for individual parameter estimates was assessed using the normal approximation. During the model-building steps, the team was a random variable (as a grouping variable), and each level pitch and their interaction terms were used as Level-1 variables. There were no Level-2 variables in this study.

Model 0 (random intercept model) represents a null model. In this model, the team was a random group variable. In Model 1, linear time (pre-and-post perturbation) was added as a predictor of team score and then tested, which improved the model in comparison to Model 0: $\chi 2$ (1) = 9.88, $p = 0.002$). In Model 2, Level-1 variables, the effect of each role's pitch (i.e., engineer, medic, transporter) were included in the best fitting model, and they did not improve model fit in comparison to Model 1, $\chi2$ (3) = 1.33, *p* = 0.723. In Model 3, the Level-1 variable interaction effect with linear time (pre-post) with role's pitch was added to the model, but it did not improve model fit in comparison to Model 2, $\chi2$ (3) = 5.67, *p* = 0.129. After that, in Model 4, also added quadratic terms for each role's pitch, which improved model fit relative to the previous model, $\chi2$ (3) = 12.68, *p* = .005. In Model 5, an interaction effect of linear time by quadratic terms of the pitch for each role was also added and tested based on the previous model. This effect did not improve in comparison to Model 4, $\chi2$ (3) = 0.07, *p* = 0.996. Thus, the test statistics of Model 4 are summarized in this study; see Table 2. However, fixed effect parameter estimates and their standard errors (*SE*) with *p*-values estimated using a normal approximation for the *t*-values for each model during the step-by-step model development process are still provided in Table 2.

Table 2. Results of Model Tests for Team Performance Score.



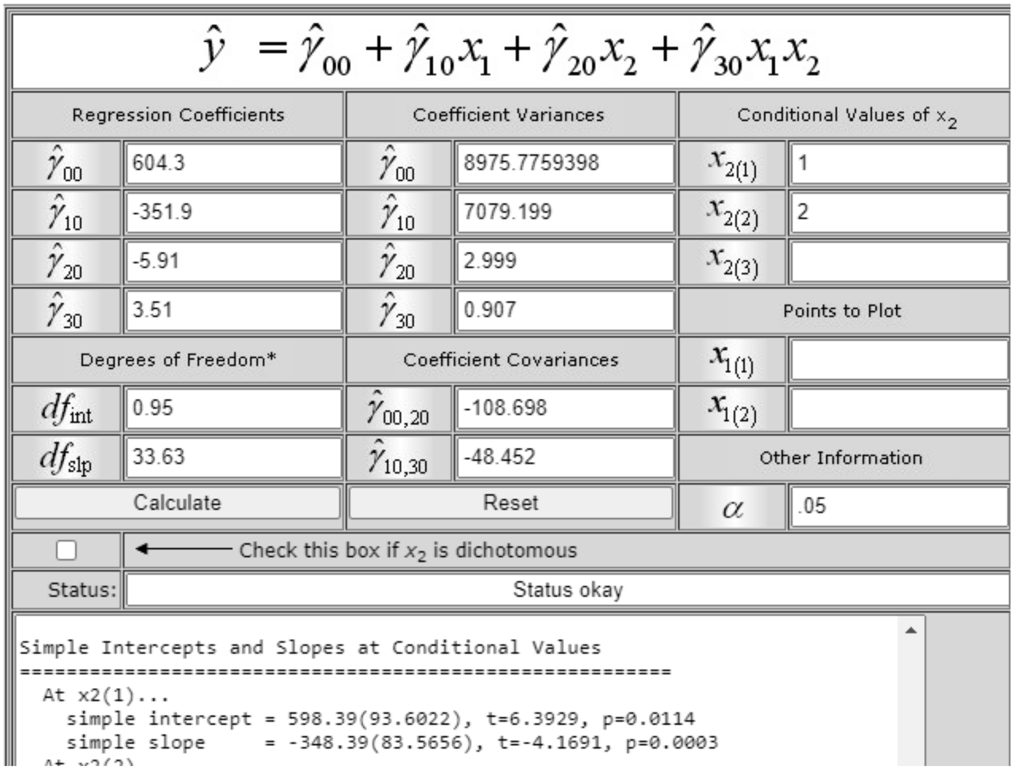
Model	# Parameter	AIC	BIC	Log Likelihood	Deviance	χ^2	χ^2 df	p-value
M0	3	673.69	679.54	-333.84	667.69			
M1	4	665.81	673.61	-328.9	657.81	9.88	1	0.002
M2	7	670.48	684.14	-328.24	656.48	1.33	3	0.723
M3	10	670.81	690.33	-325.41	650.81	5.67	3	0.129
M4	13	664.13	689.5	-319.07	638.13	12.68	3	0.005
M5	16	670.07	701.29	-319.03	638.07	0.07	3	0.996

In Model 4, the rate of change, i.e., the slope of the linear term (pre-and-post), was significantly and negatively related to team performance, 10 = -351.90, *SE* = 84.14, *p* = 0.001. The linear time indicates that team performance decreased over time, likely due to the high workload of the perturbations during the post-perturbation in Mission 2. The post-perturbation time interval had a loss of GPS signal blackout that resulted in an increase in sensitivity to differences between the high-performing and low-performing teams.

During pre-perturbation, the transporter pitch effect (Linear) was statistically significant, but negatively associated with team performance, 20 = -5.91, *SE* = 1.71, *p* = .002. It is possible that the transporter may have been scouting for victims for the medic to rescue without contributing to the team score, which leads to additional stress.

The interaction effect of the pre-and-post-perturbation (nominal variable) by pitch (Linear term) was only significant for the transporter, *t* (52) = 3.69, *p* = 0.002. According to the significant interaction effect, the simple slope calculator (table 3) was used; for the calculation steps, see Preacher et al., n.d. By using the VCOV function of the R software package, *11 x 11 Coefficient Covariances Matrix of Class* was obtained and reported in Appendix B. The asymptotic variances of the interaction effect's regression parameters were entered under "Coefficient Variances" in Table 3. The simple slope finding indicates that the *linear* transporter's pitch was negatively associated with team performance in the post-perturbation 32 = -4.169, *SE* = 1.73, *p* < 0.001. These significant findings may indicate that when the transporter gets more stressed due to roadblocks, a process reveals a relationship between pitch and team performance.

Table 3. Simple slope calculation is an interaction between two predictors within the Level 1 equation but with no predictors of these effects at Level 2.



$$\hat{y} = \hat{\gamma}_{00} + \hat{\gamma}_{10}x_1 + \hat{\gamma}_{20}x_2 + \hat{\gamma}_{30}x_1x_2$$

Regression Coefficients		Coefficient Variances		Conditional Values of x_2	
$\hat{\gamma}_{00}$	604.3	$\hat{\gamma}_{00}$	8975.7759398	$x_{2(1)}$	1
$\hat{\gamma}_{10}$	-351.9	$\hat{\gamma}_{10}$	7079.199	$x_{2(2)}$	2
$\hat{\gamma}_{20}$	-5.91	$\hat{\gamma}_{20}$	2.999	$x_{2(3)}$	
$\hat{\gamma}_{30}$	3.51	$\hat{\gamma}_{30}$	0.907	Points to Plot	
Degrees of Freedom*		Coefficient Covariances		$x_{1(1)}$	
df_{int}	0.95	$\hat{\gamma}_{00,20}$	-108.698	$x_{1(2)}$	
df_{slp}	33.63	$\hat{\gamma}_{10,30}$	-48.452	Other Information	
Calculate		Reset		α	.05

☐ Check this box if x_2 is dichotomous

Status: Status okay

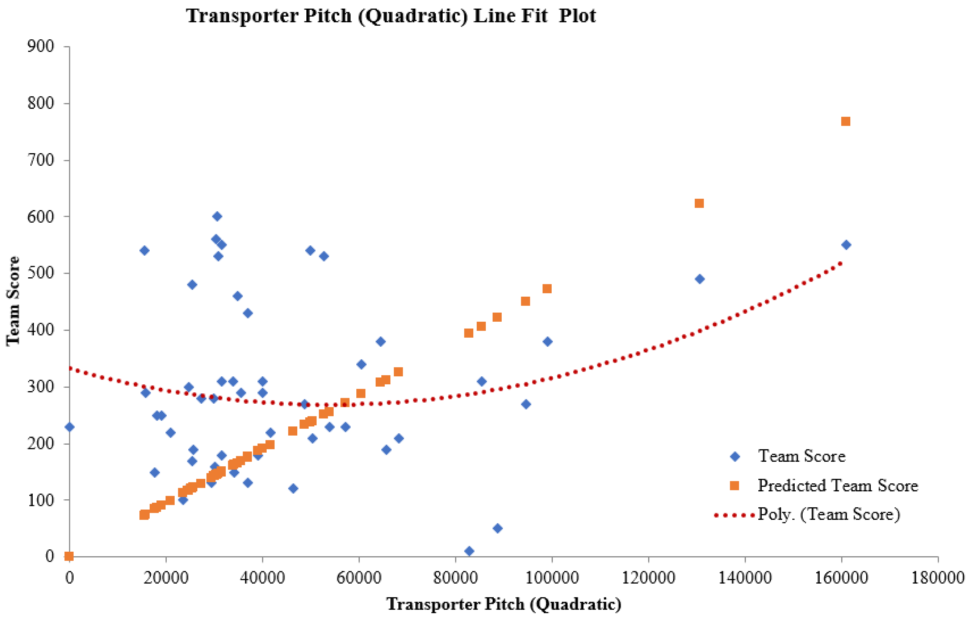
```
Simple Intercepts and Slopes at Conditional Values
=======================================================
  At x2(1)...
    simple intercept = 598.39(93.6022), t=6.3929, p=0.0114
    simple slope     = -348.39(83.5656), t=-4.1691, p=0.0003
```

On the other hand, the *quadratic* term of the transporter's pitch relationship with team performance was statistically significant and positively associated with team performance across Mission 2, 24 = 1.23, *SE* = 0.50, *p* = .002 (see Table 4). In other words, *moderate transporter pitch during a mission was associated with poorer team performance. A lower or higher transporter pitch was associated with better team performance.* The positive relationship between stress and performance can be that two team members were communicating directing to the transporter to place rescued victims in the correct subcategory zone across the task, especially during the post-perturbation duration, which means the transporter was more stressed out to recover from the roadblock to achieve the overall team goal.

Table 4. Results of Model Tests for Predicting Team Performance by Role Level Pitch across Pre-and-Post Perturbation.

	β	*SE*	*df*	*t*	*p-value*
Pre & Post	-351.90	84.14	29.58	-4.18	0.000
Engineer (Linear)	0.07	1.81	35.08	0.04	0.967
Medic (Linear)	-1.34	1.60	40.44	-0.84	0.407
Transporter (Linear)	-5.91	1.73	39.51	-3.41	0.002
Engineer (Quadric)	-0.01	0.01	35.39	-0.73	0.472
Medic (Quadric)	0.00	0.01	37.82	-0.19	0.851
Transporter (Quadratic)	0.03	0.01	38.65	3.30	0.002
Engineer (Linear) by Pre & Post	0.62	1.43	31.66	0.44	0.666
Medic (Linear) by Pre & Post	1.31	0.83	25.65	1.57	0.130
Transporter (Linear) by Pre & Post	3.51	0.95	33.63	3.69	0.001

Figure 5. Regression Plot depicting the relationship between the transporter's pitch in the quadratic term with the team score pre-and-post perturbation. The regression line shows a positive relationship between the quadratic term and the team performance score.



CHAPTER 5

DISCUSSION

The findings for RQ1 failed to reject the null hypothesis and did not meet the expectations of previous studies that found that emotional stress expected to affect performance and impact voice pitch (Kreiman & Sidtis, 2011). Although the official task experiment was not designed to cause stress to participants, it served to develop ASI, and as a result, the pre-and post-perturbation events may not have been strong enough to trigger a psychophysiological response captured in the voice.

On the other hand, the results do support the initial hypothesis of RQ2: *what portion of the team's performance is explained by each role's pitch change after a perturbation?* However, this hypothesis is only supported for the transporter task role. Considering gender was necessary because pitches differed between the sexes during the analysis. The analyzed teams had a majority of male participants (73%), followed by female participants (23%). In the teams that were analyzed, one female was assigned the role of a transporter.

A possible explanation for this significance is the task and role design; the transporter's pitch is associated with performance in a U-shaped manner. The transporter is the fastest member of the team; the transporter was likely to have higher stress when communicating in pre- and post-perturbation to gain more points. According to these results, the transporter may have the most central role in voice communication because when the transporter is under varying levels of workload and stress, and thus voice pitch has a complex relationship with performance for that role. During a task, there appears to be a U-shaped relationship between a transporter's pitch and the team's performance, with performance decreasing up to a certain point. The pitch used by transporters has a positive relationship with the team in adapting, particularly following a perturbation, as a high-level pitch has a relationship with recovery from perturbation. The analysis's initial claim suggests that pitch and stress are linked and that stress, in turn, is positively associated with performance. Therefore, by measuring pitch, it may be possible to predict levels of stress and performance.

Limitations and Future Work

The empirical results reported herein should be considered limitations. The research questions for this thesis were not originally designed for this study. There is limited evidence in survey questions, such as if the stress was experienced; as reported in the *Procedure* section, the survey contained no other questions besides "Did the team have "good emotional balance?". It will be valuable to provide more questions in the survey asking about indications of emotional stress in future studies.

For this thesis, a workload measurement is unavailable, and a given workload measurement could have influenced vocal pitch. Other considerations include that the present work differs from previous research on psychophysiology that measured vocal pitch in conjunction with heart rate variability (HRV) to indicate emotional state (Cosetl & Lopez, 2011; Urban et al., 1996). Due to the experiment being remotely administered, collecting other physiological data (e.g., heart rate) would have been challenging. However, other characteristics in a human voice, such as jitter, shimmer, pauses, and intensity, were not applied when analyzing the vocal range for differences which may have provided more insights into vocal cues during Mission 2.

In addition, there was a lack of sensitivity to cumulative team performance scores versus the number of victims rescued for that interval. For example, during the pre-and post-perturbation, depending on the team’s strategy rescuing higher-scoring victims (i.e., critical victims) versus lower-scoring victims (i.e., regular victims) may have been decided to take priority to increase team score. Another limitation is that the pre-and-post-perturbation score was not normalized based on their calculation over time; also only considered high-and-low-performing teams chose from three conditions of ASI (DOLL, CRA, CMU).

Another limitation of the experiment, not all participants were native English speakers, a potential contributor to variations in pitch. Aside from testbed technical limitations, this study also relied on remote participation, which required participants to use personal audio microphones to capture their voices, along with internet connections and computer software; all internet connections, microphones, and software were tested to meet standards for usability.

CONCLUSION

Future ASI developments could intervene if the vocal pitch of a player deviates from baseline, which may become increasingly important in understanding human teams interacting with their environment. The purpose of this thesis was to investigate the relationship between vocal pitch and team performance in simulated search and rescue missions. We conclude that vocal pitch pre- and post-perturbation is related to team performance, particularly for specific roles. This work provides a conceptual framework for the relationship between voice pitch and team performance. Examining psychophysiological responses, including voice pitch related to team performance, can inform the ongoing development of robust, intelligent ASI tools that can support dynamic, high-stress team environments.

REFERENCES

Andreassi, J. L. (2013). *Psychophysiology: Human Behavior & Physiological Response*. Psychology Press.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*, 614–636. https://doi.org/10.1037/0022-3514.70.3.614

Brenner, M., Doherty, E. T., & Shipp, T. (1994). Speech measures indicating workload demand. *Aviation, Space, and Environmental Medicine*, *65*, 21–26.

Bunn, J. C., & Mead, J. (1971a). Control of ventilation during speech. *Journal of Applied Physiology*. https://doi.org/10.1152/jappl.1971.31.6.870

Bunn, J. C., & Mead, J. (1971b). Control of ventilation during speech. *Journal of Applied Physiology*. https://doi.org/10.1152/jappl.1971.31.6.870

Cabrera Cosetl, R., & Baez Lopez, J. M. D. (2011). Voice Stress Detection: A method for stress analysis detecting fluctuations on Lippold microtremor spectrum using FFT. *CONIELECOMP 2011, 21st International Conference on Electrical Communications and Computers*, 184–189. https://doi.org/10.1109/CONIELECOMP.2011.5749357

Campbell, J., & Ehlert, U. (2012). Acute psychosocial stress: Does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology*, *37*(8), 1111–1134. https://doi.org/10.1016/j.psyneuen.2011.12.010

Chakraborti, T., Sreedharan, S., & Kambhampati, S. (2017). Balancing Explicability and Explanation in Human-Aware Planning. *ArXiv E-Prints*, *1708*, arXiv:1708.00543.

Committee on the Science of Team Science, Board on Behavioral, Cognitive, and Sensory Sciences, Division of Behavioral and Social Sciences and Education, & National Research Council. (2015). *Enhancing the Effectiveness of Team Science* (N. J. Cooke & M. L. Hilton, Eds.). National Academies Press (US). http://www.ncbi.nlm.nih.gov/books/NBK310387/

Corral, C. C., Tatapudi, K. S., Buchanan, V., Huang, L., & Cooke, N. J. (2021). Building a Synthetic Task Environment to Support Artificial Social Intelligence Research. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *65*(1), 660–664. https://doi.org/10.1177/1071181321651354a

Crosswell, A. D., & Lockwood, K. G. (2020). Best practices for stress measurement: How to measure psychological stress in health research. *Health Psychology Open*, *7*(2), 2055102920933072. https://doi.org/10.1177/2055102920933072

Demir, M., Likens, A. D., Cooke, N. J., Amazeen, P. G., & McNeese, N. J. (2019). Team Coordination and Effectiveness in Human-Autonomy Teaming. *IEEE Transactions on Human-Machine Systems*, 1–10. https://doi.org/10.1109/THMS.2018.2877482

Demir, M., McNeese, N. J., & Cooke, N. J. (2017). Team situation awareness within the context of human-autonomy teaming. *Cognitive Systems Research*, *46*, 3–12. https://doi.org/10.1016/j.cogsys.2016.11.003

Fairbanks, G., & Pronovost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotions. *Speech Monographs*, *6*, 87–104. https://doi.org/10.1080/03637753909374863

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(45), 17538–17542. https://doi.org/10.1073/pnas.0805903105

Freeman, J., Huang, L., Wood, M., & Cauffman, S. J. (2022). Evaluating Artificial Social Intelligence in an Urban Search and Rescue Task Environment. In N. Gurney & G. Sukthankar (Eds.), *Computational Theory of Mind for Human-Machine Teams* (pp. 72–84). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-21671-8_5

Frith, C. D., & Frith, U. (2006). How we predict what other people are going to do. *Brain Research*, *1079*(1), 36–46. https://doi.org/10.1016/j.brainres.2005.12.126

Gage, N., & Baars, B. (2018). *Fundamentals of Cognitive Neuroscience—2nd Edition*. https://www.elsevier.com/books/fundamentals-of-cognitive-neuroscience/gage/978-0-12-803813-0

Gaillard, A. W. (1993). Comparing the concepts of mental load and stress. *Ergonomics*, *36*(9), 991–1005. https://doi.org/10.1080/00140139308967972

Google Meet. (n.d.). Retrieved December 13, 2022, from https://meet.google.com/?pli=1

Gorman, J. C., Grimm, D. A., Stevens, R. H., Galloway, T., Willemsen-Dunlap, A. M., & Halpin, D. J. (2019). Measuring Real-Time Team Cognition During Team Training. *Human Factors*, 0018720819852791. https://doi.org/10.1177/0018720819852791

Hart, S., & Hauser, J. R. (1987). Inflight application of three pilot workload measurement techniques. *Aviation, Space, and Environmental Medicine*. https://www.semanticscholar.org/paper/Inflight-application-of-three-pilot-workload-Hart-Hauser/aac311e8da1a8982d7fca5ecb02bf6402b0dd211

Jerčić, P., Sennersten, C., & Lindley, C. (2020). Modeling cognitive load and physiological arousal through pupil diameter and heart rate. *Multimedia Tools and Applications*, *79*(5), 3145–3159. https://doi.org/10.1007/s11042-018-6518-z

Khawaja, M. A., Chen, F., & Marcus, N. (2012). Analysis of Collaborative Communication for Linguistic Cues of Cognitive Load. *Human Factors*, *54*(4), 518–529. https://doi.org/10.1177/0018720811431258

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the Effectiveness of Work Groups and Teams. *Psychological Science in the Public Interest*, *7*(3), 77–124. https://doi.org/10.1111/j.1529-1006.2006.00030.x

Kreiman, J., & Sidtis, D. (2011). *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Wiley-Blackwell. https://doi.org/10.1002/9781444395068

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming With a Synthetic Teammate: Insights into Human-Autonomy Teaming. *Human Factors*, *60*(2), 262–273. https://doi.org/10.1177/0018720817743223

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R* (1 edition). Chapman and Hall/CRC.

Murphy, R. R., Tadokoro, S., Nardi, D., Jacoff, A., Fiorini, P., Choset, H., & Erkmen, A. M. (2008). Search and Rescue Robotics. In B. Siciliano & O. Khatib (Eds.), *Springer Handbook of Robotics* (pp. 1151–1173). Springer. https://doi.org/10.1007/978-3-540-30301-5_51

OpenSMILE 3.0—AudEERING. (n.d.). Retrieved December 13, 2022, from https://www.audeering.com/research/opensmile/

Palada, H., Neal, A., Tay, R., & Heathcote, A. (2018). Understanding the causes of adapting, and failing to adapt, to time pressure in a complex multistimulus environment. *Journal of Experimental Psychology: Applied*, *24*, 380–399. https://doi.org/10.1037/xap0000176

Pisanski, K., Groyecka-Bernard, A., & Sorokowski, P. (2021). Human voice pitch measures are robust across a variety of speech recordings: Methodological and theoretical implications. *Biology Letters*, *17*(9), 20210356. https://doi.org/10.1098/rsbl.2021.0356

Pisanski, K., Nowak, J., & Sorokowski, P. (2016). Individual differences in cortisol stress response predict increases in voice pitch during exam stress. *Physiology & Behavior*, *163*, 234–238. https://doi.org/10.1016/j.physbeh.2016.05.018

Preacher, K. J., Curran, P. J., & Bauer, D. J. (n.d.). *Simple intercepts, simple slopes, and regions of significance in HLM 2-way interactions*. Quantiative Psychology. Retrieved January 8, 2017, from http://www.quantpsy.org/interact/hlm2.htm

Re, D. E., O'Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for Very Low and Very High Voice Pitch in Humans. *PLOS ONE*, *7*(3), e32719. https://doi.org/10.1371/journal.pone.0032719

Roßnagel, C. (2000). Cognitive load and perspective-taking: Applying the automatic-controlled distinction to verbal communication. *European Journal of Social Psychology*, *30*, 429–445. https://doi.org/10.1002/(SICI)1099-0992(200005/06)30:3<429::AID-EJSP3>3.0.CO;2-V

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On Teams, Teamwork, and Team Performance: Discoveries and Developments. *Human Factors*, *50*(3), 540–547. https://doi.org/10.1518/001872008X288457

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training: Robert W. Swezey, Eduardo Salas: Books. In R. W. Swezey & E. Salas (Eds.), *Teams: Their Training and Performance* (pp. 3–29). Ablex Pub. Corp. http://www.amazon.com/Teams-Training-Performance-Robert-Swezey/dp/089391942X

Scherer, K. R. (1986). Voice, Stress, and Emotion. In M. H. Appley & R. Trumbull (Eds.), *Dynamics of Stress: Physiological, Psychological and Social Perspectives* (pp. 157–179). Springer US. https://doi.org/10.1007/978-1-4684-5122-1_9

Sobol-Shikler, T. (n.d.). *Analysis of affective expression in speech*.

Sondhi, S., Khan, M., Vijay, R., & K. Salhan, A. (2015). Vocal Indicators of Emotional Stress. *International Journal of Computer Applications*, *122*(15), 38–43. https://doi.org/10.5120/21780-5056

SPSS Statistics | IBM. (n.d.). Retrieved February 16, 2023, from https://www.ibm.com/products/spss-statistics

Story, B., Titze, I., & Hoffman, E. (2018). *Vocal tract area functions for an adult female speaker based on volumetric imaging: The Journal of the Acoustical Society of America: Vol 104, No 1*. https://asa.scitation.org/doi/10.1121/1.423298

Stress, n.: Oxford English Dictionary. (n.d.). Retrieved March 13, 2023, from https://www.oed.com/viewdictionaryentry/Entry/191511;jsessionid=451F2E4A66EE246334128A71D0D2C3EC

Teigen, K. H. (1994). Yerkes-Dodson: A Law for all Seasons. *Theory & Psychology*, *4*(4), 525–547. https://doi.org/10.1177/0959354394044004

Thayer, J. F., & Lane, R. D. (2009). Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience & Biobehavioral Reviews*, *33*(2), 81–88. https://doi.org/10.1016/j.neubiorev.2008.08.004

Thornton, F., Dollinger, M., Kniesburges, S., Berry, D., Alexiou, C., & Schutzenberger, A. (2019). *Impact of Subharmonic and Aperiodic Laryngeal Dynamics on the Phonatory Process Analyzed in Ex Vivo Rabbit Models*. https://www.mdpi.com/2076-3417/9/9/1963

Tolkmitt, F. J., & Scherer, K. R. (1986). Effect of experimentally induced stress on vocal parameters. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 302–313. https://doi.org/10.1037/0096-1523.12.3.302

Urban, J. M., Bowers, C. A., Monday, S. D., & Morgan Jr., B. B. (1995). Workload, Team Structure, and Communication in Team Performance. *Military Psychology*, *7*(2), 123–139. https://doi.org/10.1207/s15327876mp0702_6

Urban, J. M., Weaver, J. L., Bowers, C. A., & Rhodenizer, L. (1996). Effects of Workload and Structure on Team Processes and Performance: Implications for Complex Team Decision Making. *Human Factors*, *38*(2), 300–310. https://doi.org/10.1177/001872089606380210

Van Puyvelde, M., Neyt, X., McGlone, F., & Pattyn, N. (2018a). Voice Stress Analysis: A New Framework for Voice and Effort in Human Performance. *Frontiers in Psychology*, *9*. https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01994

Van Puyvelde, M., Neyt, X., McGlone, F., & Pattyn, N. (2018b). Voice Stress Analysis: A New Framework for Voice and Effort in Human Performance. *Frontiers in Psychology*, *9*, 1994. https://doi.org/10.3389/fpsyg.2018.01994

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance Requires Hard Mental Work and Is Stressful. *Human Factors*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152

Williams, J., Fiore, S. M., & Jentsch, F. (2022). Supporting Artificial Social Intelligence With Theory of Mind. *Frontiers in Artificial Intelligence*, *5*. https://www.frontiersin.org/articles/10.3389/frai.2022.750763

Zavala, A. M., Day, G. E., Plummer, D., & Bamford-Wade, A. (2018). Decision-making under pressure: Medical errors in uncertain and dynamic environments. *Australian Health Review: A Publication of the Australian Hospital Association*, *42*(4), 395–402. https://doi.org/10.1071/AH16088

Zhang, Z. (2016). Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *The Journal of the Acoustical Society of America*, *139*(4), 1493. https://doi.org/10.1121/1.4944754

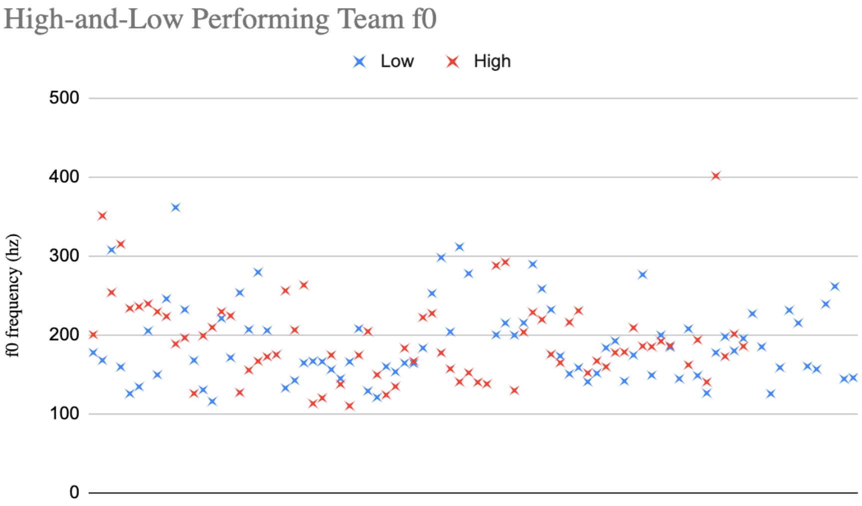
Zoom Chat. (n.d.). Zoom. Retrieved March 19, 2021, from https://zoom.us/

14.2—The General m-Stage Nested Design | STAT 503. (n.d.). PennState: Statistics Online Courses. Retrieved February 16, 2023, from https://online.stat.psu.edu/stat503/lesson/14/14.2

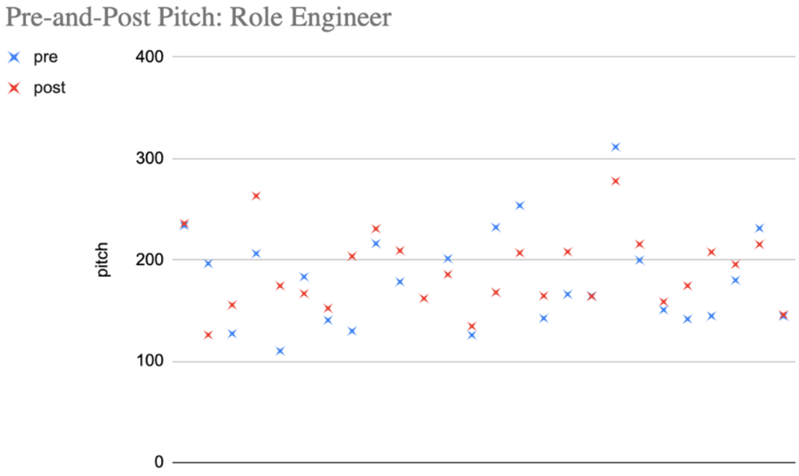
APPENDIX A

SCATTER PLOTS

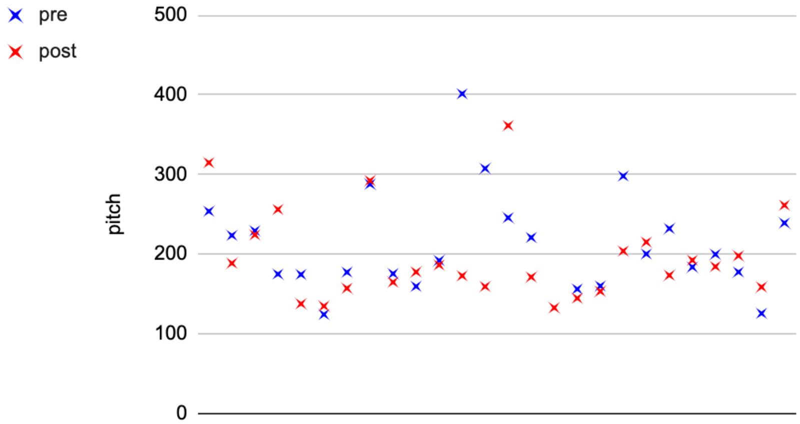
1. This chart illustrates the differences between routine and novel tasks in terms of voice pitch frequencies from 100 to 410 *Hz* between high and low-performing teams.



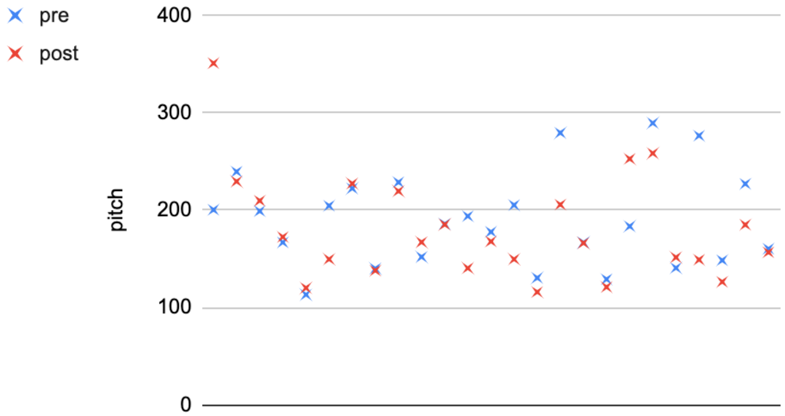
2. A scatter plot of pre-and-post perturbation voice pitch comparison in each role, i.e., medic, transporter, and engineer, across high-and-low-performing teams in mission.



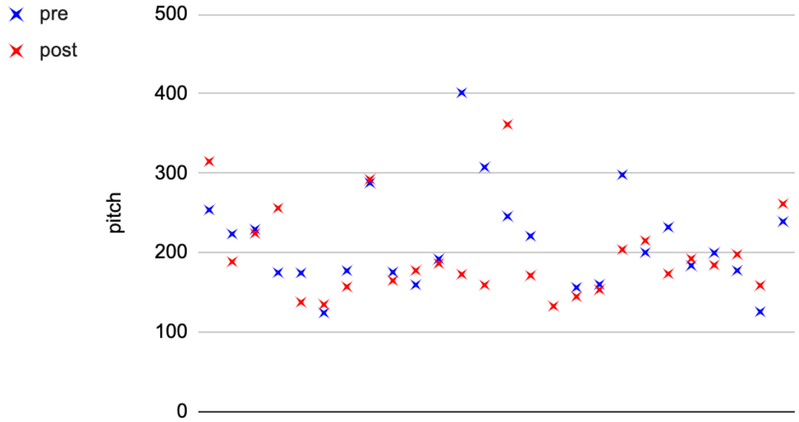
Pre-and Post Pitch: Role Transporter



Pre-and-Post Pitch: Role Medic

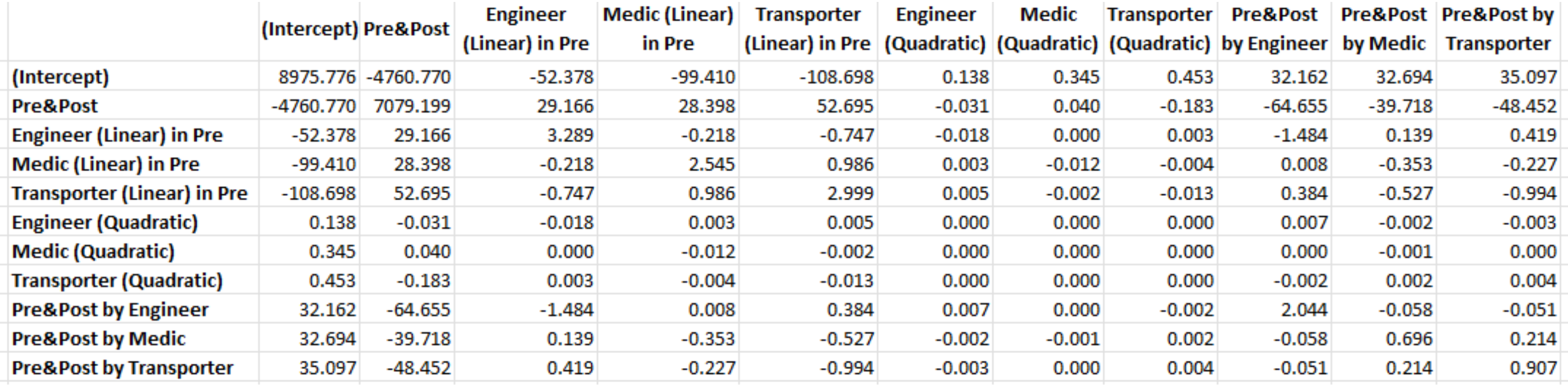


Pre-and Post Pitch: Role Transporter



APPENDIX B

11 X 11 COEFFICIENT COVARIANCES MATRIX OF CLASS



	(Intercept)	Pre&Post	Engineer (Linear) in Pre	Medic (Linear) in Pre	Transporter (Linear) in Pre	Engineer (Quadratic)	Medic (Quadratic)	Transporter (Quadratic)	Pre&Post by Engineer	Pre&Post by Medic	Pre&Post by Transporter
(Intercept)	8975.776	-4760.770	-52.378	-99.410	-108.698	0.138	0.345	0.453	32.162	32.694	35.097
Pre&Post	-4760.770	7079.199	29.166	28.398	52.695	-0.031	0.040	-0.183	-64.655	-39.718	-48.452
Engineer (Linear) in Pre	-52.378	29.166	3.289	-0.218	-0.747	-0.018	0.000	0.003	-1.484	0.139	0.419
Medic (Linear) in Pre	-99.410	28.398	-0.218	2.545	0.986	0.003	-0.012	-0.004	0.008	-0.353	-0.227
Transporter (Linear) in Pre	-108.698	52.695	-0.747	0.986	2.999	0.005	-0.002	-0.013	0.384	-0.527	-0.994
Engineer (Quadratic)	0.138	-0.031	-0.018	0.003	0.005	0.000	0.000	0.000	0.007	-0.002	-0.003
Medic (Quadratic)	0.345	0.040	0.000	-0.012	-0.002	0.000	0.000	0.000	0.000	-0.001	0.000
Transporter (Quadratic)	0.453	-0.183	0.003	-0.004	-0.013	0.000	0.000	0.000	-0.002	0.002	0.004
Pre&Post by Engineer	32.162	-64.655	-1.484	0.008	0.384	0.007	0.000	-0.002	2.044	-0.058	-0.051
Pre&Post by Medic	32.694	-39.718	0.139	-0.353	-0.527	-0.002	-0.001	0.002	-0.058	0.696	0.214
Pre&Post by Transporter	35.097	-48.452	0.419	-0.227	-0.994	-0.003	0.000	0.004	-0.051	0.214	0.907

BIOGRAPHICAL SKETCH

Jeska Clark was born in Philadelphia, Pennsylvania. After she graduated from Lakeside High School, he attended the Hugh Downs School of Communication at Arizona State University, majoring in communication in 2010. Upon graduation in 2014, She started working as a consultant in the healthcare and aerospace industries. During this time, she published: “Medical Guidelines for Commercial Suborbital Human Spaceflight: A Review” The research resulted in the *Journal of Acta Astronautica* in 2021. Ms. Clark worked as a graduate research assistant during her graduate studies.