Human Performance and Cognitive Workload in Haptic, Audio and Visual Environments

by

Mortaja AlQassab

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The ability to efficiently perform a task in a human-in-the-loop system and in multi-sensory virtual environments is highly dependent on the type of sensory feedback the operator is receiving and the amount of workload the operator is exposed to. Despite the vast amount of research on Collaborative Virtual Environments (CVEs) and Human Machine Interactions (HMIs), little is known about what type of feedback increases the performance of a human operator and what type of sensory feedback minimizes the amount of workload the operator is exposed to. While individual differences influence human performance outcomes, the physiological processes a human being set the fundamental guidelines for assessing human performance.

The purpose of this study is to evaluate the performance of participants for a combination of sensory two feedback modes (audio-visual, haptic-visual or audio-haptic) in a primary task to find the optimum feedback model for CVE and HMI applications. A concurrent secondary task is also designed to evaluate workload of each feedback mode (audio, haptic or visual) and the effect of different levels of workload on task completion time and task accuracy. For example, a car driver performs a primary task by steering the car in the correct direction. A secondary task, in the same context, would be monitoring the fuel level or checking the speed limit.

In the primary task, participants are required to press a virtual button from a set of three (right button, left button or up button). The secondary task evaluates the amount of workload the participant is exposed to in three different feedback modes (haptic, audio or visual). Each participant is required to recognize a Morse code. In this study, participants perform three trials. In first trial, participants perform one task the primary task alone. In the second and third trials, participants perform the primary task and the secondary task concurrently. The primary task evaluates human performance and includes combined sensory modalities as a feedback mode (audio-visual, haptic-visual or audio-haptic).

The time it takes the participant to press the virtual button (primary task response time), the number of correct button presses (primary task accuracy), the time it takes the participant to recognize the Morse code (secondary task response time) and the number of the correct codes (secondary task accuracy) are all collected. In addition, NASA Task Load Index (TLX) questionnaire is used after each trial to assess the subjective performance and subjective workload of participants. The data collected is tested for normality using Lilliefors test, filtered using Grubb's test to eliminate outlying data and analyzed using one-way ANOVA and multiple two-sample t-tests. A Tukey HSD is also used to show the differences between experimental conditions.

The results of this study indicate that the hypothesis that all combinations of feedback provide the same performance can be rejected for the primary task response time. For instance, the results show that the there is a difference in response time between the audio-haptic and the audio-visual feedback modes in the first, second and third trials. The results of this study also indicate that the hypothesis that all sensory feedback modes provide the same workload can be rejected for the secondary task accuracy. Results show that there is a difference between haptic and auditory conditions and shows that visual condition has a lower accuracy than the other feedback modes.

Acknowledgement

I would like to express my deepest gratitude to my advisor, Professor David Wang, for offering a continuous encouragement and support during my graduate studies at the University of Waterloo. His advice, guidance and knowledge throughout my journey from the process of finding an appropriate subject to the process of writing this thesis.

I must also express my appreciation to my readers, Professor Dana Kulić and Professor James Tung, for reading my thesis and for their constructive feedback.

Sincere thanks to my friends, locally and globally, for their endless encouragement and to my colleagues at the University of Waterloo, especially Adam Gomez and Orion Bruckman for sharing their experiences and for providing the technical support.

I am also thankful to my participants whose names cannot be disclosed in this thesis for taking part in this study. I want to acknowledge their help in collecting information. Their participation helped me completing this thesis.

I am grateful and thankful to my parents whose lifetime support, unlimited inspiration and endless encouragement made this opportunity possible. Many thanks to my brothers and sisters. Finally, special thanks to my wife for her love, patience and motivation.

Dedication

Dedicated to my mother.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and Applications

Communication is linked to different sensory modalities. In-person communication usually includes vision, hearing, touch, smell and taste. However, in Collaborative Virtual Environments (CVEs) and Human-Machine Interactions (HMIs), communication is normally limited to audio-visual communication. Current research suggests that including more channels of communication such as haptic feedback could, to some extent, enhance performance, increase the ability of an individual or a group to accomplish a task, or increase tele-presence. Tele-presence can be defined as the extent to which people perceive if they are situated in the remote location where the simulation is taking place [1]. Other studies argue that integrating another sensory modality to the communication model could convey information for operators whose other senses are preoccupied in a demanding task. Even though including more sensory modalities to CVEs and HMIs might increase telepresence[1]–[5] or enhance user experience[6]–[9], human performance and cognitive workload could degrade due to the anatomical limitation of the human sensory modalities and the cognitive limitation of the human information processing.

A primary task can be thought of as a physical task a human operator performs and a secondary task is a cognitive task where the human operator receives and processes data. In this thesis, participants are required to press a virtual button while receive Morse codes that participants should recognize. For example, a locomotive engineer's primary task is to control acceleration of the locomotive and a secondary task is to receive and interpret different railway signals. As another example, a surgeon's primary task is to operate on a human's body while monitoring his physiological signs to know any of his critical conditions. In this thesis, an experiment is designed and conducted to evaluate the performance of participants for a combination of two feedback modes in a primary task to find the optimum feedback model for CVEs and Human-Machine Interaction applications. The combined sensory modalities for the primary task are audio-visual, haptic-visual or audio-haptic. A secondary task is also designed to evaluate workload of each feedback mode and the effect of different levels of workload on task completion time and task accuracy. The sensory modalities the participants are exposed to in the secondary task are audio, haptic or visual.

1.1.1 Human-Machine Interaction (HMI)

Human-Machine Interaction (HMI) is becoming an integral part in the process of designing and optimizing systems that involve human in the loop. Because machines and software mediate most systems, HMI is crucial for human-in-the-loop system in order to provide a safe, efficient and reliable system [10].

Aviation manufacturers suggest that added haptic force feedback could convey information to people preoccupied by other auditory or visual tasks. Figure 1.1 shows a cockpit with different kinds of displays. Aviation companies have added haptic interface as a new kind of display in airplanes to convey crucial information about the aircraft status and more specifically to warn pilots about a near-stall condition [11]. Moreover, car manufacturers have recently incorporated haptic feedback into driver's seats to warn them about deviations from roads [12].

It is crucial to understand the behavior of human performance in human-in-the-loop systems. Thus, it is necessary to recognize multi-sensory feedback mode that increases the performance of a user in multitasking applications. Additionally, understanding the amount of workload operators can tolerate without interrupting the primary task is vital for the effective completion of the task.

Figure 1.1: Stick shakers are added to airplanes` control columns [11].

1.1.2 Collaborative Virtual Environments (CVEs)

Collaboration (or cooperation) is the processes of integrating the inputs of two or more users to change the state of a system. Cooperation is used more for tasks that involve simultaneous actions from users to change the characteristics such as color, position and orientation of an object [13]. Both collaboration and cooperation refers to the concurrent manipulation of an object to change its state and is used interchangeably for the purpose of this thesis.

It is important to study individual interactions of performance in CVEs. Individual interaction with the environment of each agent (the user) in CVEs can affect the holistic performance of the task. Collaboration in Virtual Environments is defined in three different levels[14]: (i) the capacity for two or more users to exist in CVEs, (ii) the ability for one user to change the state of an object in the CVE and (iii) the ability for two or more users to simultaneously change the state of an object in the CVE. Subsequently, the third level of collaboration is further categorized into two sub-levels. Users can simultaneously change two different independent properties of one object that does not cause any conflict, or users can simultaneously change the same property of an object that might cause a conflict. For instance, users can collaborate to build a software project using a versioning system like SVN [15] by modifying different parts of the project at the same time, but this collaboration may cause some conflicts. Thus, users in collaboration ought to have a common ground of understanding and a method of feedback to resolve conflicts. In this thesis, the second level of collaboration is investigated in terms of user performance and workload to find the optimum combination of sensory feedback for Collaborative Virtual Environments (CVEs).

Significant number of studies have evaluated human performance, awareness, user experience and presence in CVEs with added haptic feedback in manipulation, selection, hand-over and training tasks [1], [5], [9], [16]–[22]. One study [1] investigated if adding haptic feedback would change the characteristics of objective and subjective task performance, social presence and virtual presence in manipulation and hand-over task (Figure 1.2). The results of the aforementioned study have shown that haptic feedback significantly decreased the total completion time of the task. Results has also shown an increase in perceived performance with the added force feedback. Similar to [1], an experiment [9] was conducted to evaluate the perceived safety of users when handling and handing over objects in a CVE. In this study, the effects of visual and haptic cues on the perceived characteristic of objects such as weight, smoothness and roughness. The difference between controlled condition, where participants freely manipulate objects, and un uncontrolled condition, where participants are given additional feedback and prevented from losing an object, was investigated. The results showed that additional feedback was not significant. This might be attributable to the fact that objects behaved unrealistically as reported by the study. Another study [23] investigated the ability of a user to communicate using a haptic link in a collaborative Computer Aided Software Engineering (CASE) design task. The influence of haptic feedback in the effectiveness of training was also studied in a writing task [16] and in surgical training [19] using What You Feel Is What I Do (WYFIWID) approach where haptic force feedback is reproduced for the one agent in the CVE. Results from [16] showed that the correct responses of subjects are almost 100% correct to the expert movements. The study also indicated that haptic guidance increased user's satisfaction.

However, few studies have been conducted to investigate the effects of different sensory modalities on task performance. These studies have either evaluated the effects of different single-sensory modality feedback modes on task performance or a two-sensory-modality combinations against a single-modality feedback mode. Moreover, none of these studies have investigated the effects of cognitive workload in human performance for different sensory feedback modes. For instance, one study [24] has investigated auditory and haptic feedback modes in a collaborative environment to guide blind people for referencing applications. The results of this study shows that haptic feedback can be used as an additional communication as it conveys much more information. A second study [25] evaluated human performance for verbal feedback mode, haptic feedback mode and verbal-haptic feedback mode in a collaborative pointing tasks. In the latter study, the results indicate that participants exposed to Visual (V) and Haptic-Visual combined had a relatively reduced response time to complete the pointing task. While most the studies in the literature suggested the use of an additional sensory modality (haptic feedback), these studies evaluate specific application such as surgical simulators and CASE. Those studies only evaluated two conditions (with/without haptic feedback) and did not evaluate performance with the presence of another cognitive task.

In this thesis, the relative performance between two- sensory modality feedback mode combinations, namely, Visual-Audio (VA), Visual-Haptic (VH) and Audio-Haptic (AH), is evaluated. In addition, cognitive workload is evaluated for three single-modality feedback modes (Visual (V), Audio (A) and Haptic (H)).

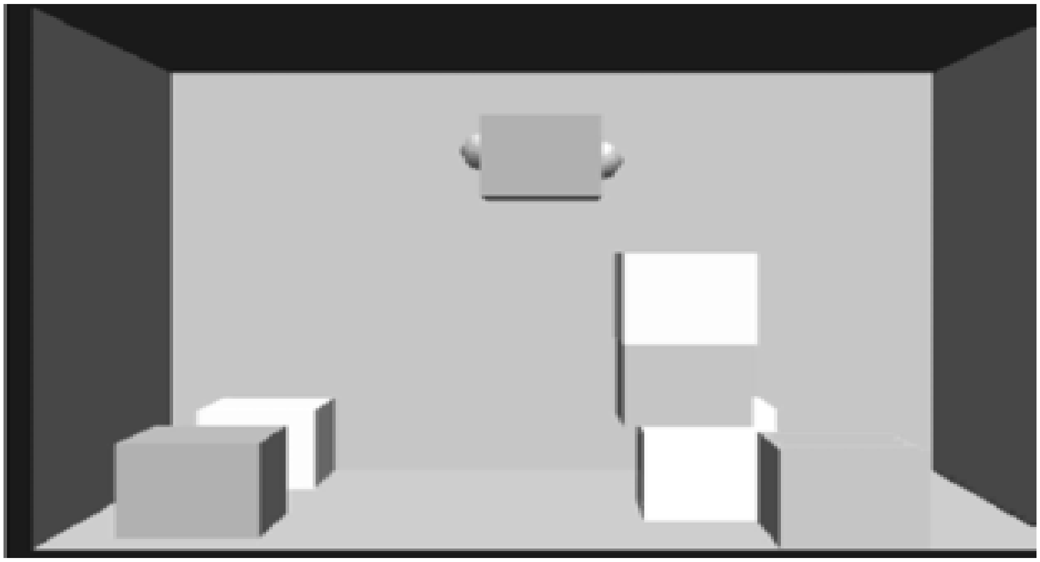


Figure 1.2: A CVE of object handling and hand-over task [1].

1.2 Thesis Outline

This thesis is organized as follows:

Chapter 2 introduces fundamental concepts of visual, auditory and haptic perception. This chapter also explains, in detail, human information capacity and examines the limitations of human cognitive ability. A survey of Collaborative Virtual Environments (CVEs) and relevant research are also presented.

Chapter 3 presents objective and subjective human performance measures for simple button-pressing task in three conditions: visual-audio, visual-haptic and haptic-audio. Performance measures consists of the efficiency and accuracy of participants to accomplish the task. This chapter also analyzes the objective measures of cognitive workload and subjective workload in a secondary task. Subjective measures of performance and workload collected using NASA Task Load Index questionnaire of workload are also presented. Subsequently, these measures are analyzed using Analysis of Variance (ANOVA) to determine the optimum combination of modalities. In this thesis, I present the study carried out to evaluate the proposed hypotheses:

- $H1_{null}$: All combinations of feedback modalities (Audio-Visual, Haptic-Audio or Haptic-Visual) the same performance and workload on users in primary task.
- $H1_{alternative}$: A certain combination of feedback modalities (Audio-Visual, Haptic-Audio or Haptic-Visual) increases performance and workload with primary task.
- $H2_{null}$: All combinations of feedback modalities (Audio, Haptic or Visual) impose the same amount of cognitive workload on users.
- $H2_{alternative}$: A specific feedback modality (Audio, Haptic or Visual) imposes more workload on users than other feedback modalities.
- $H3a_{null}$: There are not differences in means in effect of condition variation.
- $H3a_{alternative}$: A specific feedback modality (Audio-Visual, Haptic-Audio or Haptic-Visual) has a different performance due to condition variation.
- $H3b_{null}$: There are not differences in means of different feedback modalities in effect of the secondary task.
- $H3b_{alternative}$: A specific feedback modality (Audio-Visual, Haptic-Audio or Haptic-Visual) has a different performance due to the secondary task.
- $H3c_{null}$: There is not interaction between condition variation and the secondary task.
- $H3c_{alternative}$: There is an interaction between the secondary task and condition variation with at least one conditions.

Chapter 4 assesses the effects of the secondary task on the primary task measures (response time and accuracy) using a two-way ANOVA between the first trial and the second trial.

Finally, Chapter 5 summarizes the study results, gives concluding remarks on the findings presented in Chapter 3 and Chapter 4 and presents future research.

Chapter 2

Background

2.1 Introduction

Human beings working on HMIs often multitask by reacting to stimuli from different sensory modalities. The limitations inherent to the human sensory modalities, however, can widely affect the task performance. This chapter provides the background knowledge on perception processes of the visual sensory modality, the auditory sensory modality and haptic sensory modality. This chapter also explains limitations of the human information capacity from different perspective. It also looks at the related work that has been done in the area of Haptic, Audio and Visual Environments (HAVEs).

2.2 Visual Perception

Photoreceptors are the fundamental building blocks for visual perception. There are five types of photoreceptors, and they change electric potential to a change in light intensity and frequency. There are two types of photoreceptors, rods and cones. Cones are used for daylight vision because they are less sensitive to light while rods are used for night vision because of their high sensitivity to light. Cones are specialized in color vision and rods are only suited for black and white vision [26]. The aforementioned study has shown that vision is the most dominant sensory modality.

Visual perception is the building of a mental model of physical objects reflected by visible light. Visual recognition for novel forms or objects requires a reasoning process by deducing their function. This process is called categorization [27]. There are many factors that affect our visual perception. Some of these factors are the object size, object location and object orientation. The absence of a distinct horizon can change the apparent size of an object and might cause a visual illusion [28]. Furthermore, visual perception is dependent on the processing of multiple views of an object; thus, the process is reliant on the object orientation [29]. Most importantly, object recognition time is a factor of the knowledge of the object and its location on viewpoint [28].

Human Visual Limitations

The human visual system has some limitations. These include the human field of view, visual precision and the visible spectrum of light [30]. Because of the position and orientation of our eyes, our eyes can only perceive scenes that span 200 degrees horizontal and 120 degrees vertical. The eye field of view generally degrades with age. In addition, the visual precision, also called angular resolution, which is defined as smallest object the human eye can capture is limited to 60 arc seconds [31]. The visible spectrum of light is another limitation of the human visual system. The frequencies of light that the human eye can see 430 to 770 Tera Hertz.

2.3 Auditory Perception

Auditory perception is the process of detecting pressure differential from the surroundings, converting it to impulses and sending it to the brain. Sound has two fundamental features, pitch, loudness [30]. Sound pitch which characterized by the frequency of the sound waves. Sound loudness is reflected by the intensity.

The ear consists of the inner and outer ears. The outer ear is made up of cartilage, which is a rigid skin and is also called the pinna. The ear pinna has crucial functions. It collects most of the sound energy, acts as an amplifier of sound waves [30]. Thus, it increases the sensitivity of the ear. A canal connects the outer ear to the eardrum in the inner ear. The eardrum is connected to miniature bones that vibrate when there is a pressure differential. Vibrations are conducted through the miniature bones and are converted to an electrical potential via the ear nerves.

Human Auditory Limits

Sound sources convey a lot of information through the harmonics of the sound waves. However, people's auditory system is limited in several ways. The human location can limit the ability of humans to perceive sounds as sound waves decay very fast following the inverse square law. Another limit of the human audition is the ranges of frequencies humans can perceive. Unlike other creatures, humans can only perceive frequencies in the range of 20 Hz to 20 kHz. Humans are also limited to sound intensities in the range of 0 to 120 decibels.

2.4 Haptic Perception

The comprehension of the sense of touch has not yet been concrete. Nevertheless, the physiology and anatomy of the sense of touch has been more clear post World War II when E. H. Weber systematically conducted studies on the sense of touch [32] in different parts of the body, and he illustrated that the sensitivity for which human perceive touch differs from one part of the body to another by using the current standard (the two-point threshold method).

Haptic, from the Greek word *heptesthai*, meaning the sense of touch [33], has many forms related to each of the functionalities of the somatosensory systems. The functionalities that the human haptic system perform are the sensing of temperature, pressure, vibration, joint movement and pain. All the functionalities of the system are carried out by the human haptic system receptors (Mechanoreceptors). Table 2.1 summarizes the characteristics of each of the mechanoreceptors. Haptics incorporates two types of sensing, proprioceptive and tactile. Proprioceptive and tactile haptics processes are crucial to understand because the human operator utilizes this sense in making decisions, generating expressions and emotions and providing guidance and control.

Mechanoreceptors	Type of Sensing	Depth	Rate of Adaptation	Frequency Range (Hz)
Merkel Disk	Tactile	Closer to the skin	Slowly adapting	0.3 – 3
Ruffini corpuscle	Tactile / Proprioceptive	Deeper in the skin	Slowly adapting	15 - 400
Meissner corpuscle	Tactile	Closer to the skin	Rapidly adapting	3 - 40
Pacinian corpuscle	Tactile / Proprioceptive	Deeper to the skin	Rapidly adapting	10 - 1000

Table A.1: The characteristics of the human haptic system mechanoreceptors [34].

Human Haptic Limitations

Human haptic limitations play an important role in the design of touch displays. In other words, if the precision of the proprioceptive of tactile haptics is known, the realism, quality, flexibility and speed of the haptic display can be designed to only include these limits. The human haptic system is limited in the force resolution the human can distinguish, the frequency of haptic stimuli the human can discriminate and precision of the movements the human can discriminate [34]. This is also called the Just Noticeable Differences (JNDs) of the human haptic system. The precision of the human haptic system is 0.06 N and uses maximum force the human haptic system can generate is 400 N. The frequencies the human can distinguish are in the range of 0.3 to 1000 Hz [35].

2.5 Human Performance Evaluation in Haptic, Audio and Visual Environments (HAVEs)

Haptic, Audio and Visual Environments (HAVEs) involve the reproduction of sensory cues via computer peripherals. HAVEs range from simple single-sensory environments to sophisticated multi-sensory and multi-dimensional environments. Complex HAVE systems usually consist of three modalities: haptic technology, binaural sound, and 3D visuals. Each component has different effects that should be handled to build a virtual environment that is efficient and realistic. Additionally, virtual systems usually involve various kinds of navigation and selection tasks.

There has been significant research to quantitatively and qualitatively evaluate single-modality and human performance in HAVEs. In [36], auditory sensory feedback was evaluated in a collaborative visual and haptic environment quantitatively by measuring the task completion time and also qualitatively by interviewing subjects. It has been shown that adding auditory feedback has enhanced the performance of participants in an object manipulation application. In [37], a standard Fitts` task was used to evaluate human performance in a selections task with and without haptic feedback. Researchers in [38] have evaluated the addition of visual feedback in haptic-enabled Virtual Environment (VR) in the influence of visually observing object deformation in the user perception of static and dynamic friction. Recently, haptic feedback has received a wide attention from researchers. In a motion tracking application, for example, where a vibro-tactile feedback is provided for users if they deviate from the desired trajectory, performance was studied [39]. Although the addition of haptic feedback in [39] has not reduced motion errors, it enhanced user experience reported by survey responses. Haptic feedback has also been recruited in training simulators. The ability of trainees to learn faster by adding haptic feedback to computer simulation for coursework teaching and driving training was evaluated in [40], [41] respectively. Haptic feedback has also been used as an assistance in a writing training application between an expert and a beginner [16]. In a tele-operated tasks, the added haptic feedback has improved the user`s performance to remotely manipulate objects using a robotic arm [3].

Additionally, a few studies have looked into the evaluation of dual-sensory-modality performance in HAVEs. In a collaborative environment, verbal communication and haptic feedback were evaluated in a navigation task as well as a combined haptic and verbal in terms of task completion time [25]. Results show that participants under haptic only condition take longer to complete the task. A hockey game was tested using different combinations of dual-sensory-modality and three-sensory-modality in a collaborative environment. Dual-sensory modality and three-sensory-modality combinations were also studied in [42] by presenting stimuli and the user responds by pressing the corresponding button.

Most of the current research focuses on evaluating quantitative task performance [3], [5], [7], [13], [18], [43]–[45], tele-presence and social presence [1], [7], [17], [46], subjective performance [20], [47]–[49], user experience [9], [21], [40] and subjective perceived safety [9]. In this thesis, subjective measures and objective measures are used to assess different combined sensory-modality. This thesis also evaluates cognitive workload for different single-sensory modalities and investigates the effects of workload on task performance.

2.6 Human Information Capacity

Psychological studies of human information capacity propose that human ability to process information is limited and the sensing channel in which information is transmitted is limited in bandwidth. This can be attributable to the capacity of the working memory to hold information. Information transmission does not depend on a unit of measurement (unit-less). Haptic channel, auditory channel or visual channel are all considered to be limited. The limitations of these channels are discussed in many studies [50]–[52].

A potential study that explains the limitations of human information capacity was proposed in [50]. The Fitts model measures the difficulty of a task. Fitts conducted three different experiments for tapping and object transfer tasks to study the human motor system's capacity. Fitts` original tapping experiment consisted of two rectangular plates mounted between two error plates. The distance between the two plates A and the width of the plate W were varied. The participants were required to tap each plate with a stylus without touching the error plate on each side. The stylus was connected to an electronic device which recorded hits and misses. Tapping experiments are very important because they are very common in computer interactions. As a result, they were extensively surveyed by many researchers. The second experiment was a disc transfer experiment where participants grasp a washer from one pin and place inside another pin. The third experiment was pin transfer task where participants grasp a pin from one hole and place it inside another hole. Different experimental conditions were designed by varying disc diameters, hole diameters, pin sizes and the distance between targets. Variations of *As* and *Ws* built 16 different experimental conditions. Results over all three experiments revealed that more errors were made when the width of target and the distance between the targets increased.

$$ID = \log_2\left(\frac{2A}{W}\right) \tag{2.1}$$

The Index of Difficulty proposed by Fitts is expressed in Equation (2.1). A represents the amplitude which is the distance to the center of the target and W represents the width of the target. In Fitts' original experiments, the work is based on Shannon`s theorem of information capacity [53]. Shannon`s theorem defines the channel capacity of the human motor system and is expressed in Equation (2.2).

$$C = B\log_2\left(\frac{S+N}{N}\right) \tag{2.2}$$

where channel capacity is denoted by C, signal power is denoted by S, noise is denoted by N and channel`s bandwidth is denoted by B. With this expression, Fitts derived his model of movement time to a target. Fitts developed a model that predicts the movement time to a target. The movement time is a function of the distance to the target and width of target.

$$MT = a + b\log_2\left(\frac{2A}{W}\right) \tag{2.3}$$

Here, MT is the movement time in seconds, and the logarithmic term is the Index of Difficulty (ID) and a and b are empirical constants that can be determined using regression analysis on the data. MT and ID are highly correlated; the movement time to a target increased when the Index of Difficulty increases.

George Miller [51] has conducted extensive analysis of existing psychological experiments to quantify the information capacity of humans. The limit to which human beings can hold information is called “absolute judgement”. The amount of information can be described as the number of bits per unit time or overall number of input bits. In the first case, increasing the amount of information can be accomplished by sending fixed number of bits of information in a small time interval. In the second case, increasing the amount of information can be accomplished by increasing the number of bits of information. In Millers experiment, the second case is considered.

The analysis was concerned with the retained information when a subject is exposed to a stimulus. Different stimuli were studied such as acoustic, haptic and visual. Miller analyzed both the capacity of short term memory and the capacity of the working memory. The limits of the short-term memory were referred to as “absolute judgement” and are limited in terms of the “bits” of information it can hold. The limits of the working memory were referred to as “the span of working memory” and are limited in terms of the “chunks” of information it can hold. In this thesis, the capacity of the short-term memory is considered. This is reflected in the number of bits in which the amount of workload is significantly increased in any stimulus.

Chapter 3

Human Performance and Cognitive Workload in Collaborative Haptic, Audio and Visual Environments

3.1 Introduction

Human Performance and cognitive workload analysis problems arise in many areas especially in time-dependent tasks. Human performance assessment and improvement is a crucial research area for automation, Human-Machine Interface (HMI) design, Collaborative Virtual Environments (CVE) and tele-presence. While individual differences influence human performance outcome, the physiological processes set the fundamental guidelines for assessing human performance.

Many human interactions with HMIs involve primary mission-control task as well as a more cognitively demanding secondary task. For example, in driving, the primary task could be controlling the steering to stay on the road whereas the secondary task could be reading signs along the road. An experiment is designed to evaluate human performance and cognitive workload using different combinations of sensory feedback modalities with both primary and secondary task. The experiment is divided into two components. For example, a car driver performs a primary task by steering the car in the correct direction. A secondary task would be receiving direction information from a GPS.

The experiment has two components. The first component evaluates human performance with a primary task. The results from the primary task are reported and analyzed. Human performance is measured based on objective and subjective criteria. Objective criteria includes the completion time of primary task and correctness rate (accuracy) between trials. The subjective criteria is from the NASA Task Load Index (TLX) questionnaire responses. In the primary task, participants are asked to press a virtual button from a set of three buttons. Participants receive indications from two of the three senses to press the appropriate virtual button. The virtual button to be pressed is set randomly at a fixed time interval.

The second component evaluates cognitive workload with a secondary task. The results from the secondary task are reported and analyzed using one-way ANOVA and multiple two sample t-tests. Human workload is measured based on objective and subjective criteria. Objective criteria include the completion time of secondary task and correctness rate (accuracy) between feedback modalities. In the secondary task, participants are asked to recognize a Morse code (decoded using the different feedback modalities).

The hypothesis is stated in Section 3.2 and the experimental setup is detailed in Section 3.3. The procedure is explained in Section 3.4 and the results are reported in Section 3.5 with further discussion in Section 3.6.

3.2 Hypothesis

For the primary task component, the hypothesis is:

- H_{null}: All combinations of feedback modalities (Audio-Visual, Haptic-Audio or Haptic-Visual) have the performance with respect to the response time and accuracy of the primary task (pressing a virtual button) and with respect to the subjective perceived performance from questionnaire entries.

- $H_{alternative}$: A certain combination of feedback (Audio-Visual, Haptic-Audio or Haptic-Visual) increases performance with respect to the response time and accuracy of the primary task (pressing a virtual button) and with respect to the subjective perceived performance from questionnaire entries.

For the secondary task component, the hypothesis is:

- H_{null}: All combinations of feedback modalities (Audio, Haptic or Visual) impose the same amount of workload on users with respect to quantitative (response time/accuracy) and qualitative (subjective workload) analysis.
- $H_{alternative}$: A specific feedback modality (Audio, Haptic or Visual) imposes more workload on users than other feedback modalities with respect to quantitative (response time/accuracy) and qualitative (subjective workload) analysis.

3.3 Methods

This section describes the participants who volunteered for this experiment, the equipment used, the experimental setup and procedures for the primary task.

3.3.1 Participants

Twenty-nine participants volunteered for this experiment. All participants are male and female undergraduate and graduate students from the University of Waterloo. All participants were regular computer users. Demographic data has not been collected from participants. The participants were recruited using University of Waterloo administered e-mail lists or posters.

The experiment was conducted in accordance with the University of Waterloo Ethics Regulations. Consent was acquired from each participant prior to the start of the experiment. The experiment was approved as a study involving human participants by the University of Waterloo Research Ethics (ORE# 20633).

3.3.2 Apparatus

PHANToM Omni haptic device manufactured by Geomagic (formerly SensAble Technologies Inc.) [54] , as shown in Figure 3.1, was used in this experiment. The PHANToM Omni haptic device is a 3-degree of freedom robot used to send simulated forces from a computer simulation to a user. The haptic device was used as a cursor and as feedback device simultaneously. The PHANoM Omni generate a maximum 3.3 N forces. The force used in this experiment is 1.25 N.

Figure 3.1: PHANToM Omni haptic device from Geomagic (formerly SensAble Technologies).

Figure 3.2: Bose A20 aviation noise cancelling headset.

A noise cancelling headset was used to transmit auditory feedback to participants. The A20 Aviation headset, manufactured by Bose Corporation, has been designed for pilot communication in an unpressurized piston-engine aircraft (Figure 3.2). Thus, this headset attenuates a vast range of noise frequencies. The noise cancellation helped isolate all other distrations that might affect participant responses. The PHANToM Omni haptic device, for example, generates noise when applying forces or vibrations to the user. These noises were eliminated by the headset.

The haptic device is connected to a computer through an IEEE 1394 FireWire interface card and the headset is connected through an auxiliary interface to a 3.5 mm TRS audio port. A 21-inch computer monitor was used to display the Graphical User Interface of the experiment. All apparatus is connected to a computer running a 64-bit Windows 7, a 3.50 GHz Intel i5 processor, 8 GB RAM and a 2 GB nVIDIA GeForce GTx 770 graphics card.

Unity3D Game Engine 4.6.1 was used to implement the GUI and integrate all sensory modality interfaces. Two processes, haptic and graphic, were handled by Unity3D. The haptic process was running at 1000 Hz update rate. The haptic process was integrated using the haptic plugin for Unity3D developed by [55]. Some force functions were added to the plugin to fit the functionalities needed for the experiment. The graphic process is handled by Unity3D, and it runs at 60 frames per second. The haptic device, audio headset and computer monitor were all interfaced to Unity3D as shown in Figure 3.4 and the functionalities and the logic of the experiment was written using C# scripts.

The GUI consists of three virtual buttons with a blue arrow that indicates the direction of the button as shown in Figure 3.3. The buttons are arranged in a right, up and left positions. The haptic characteristic of the buttons includes a static friction and dynamic friction. Participants feel the rough surfaces of the buttons when they touch them or move over them. The button also resizes when touched by the cone cursor to simulate a button press. A yellow cone cursor was designed to navigate the environment. Audio sensory feedback is conveyed to the participants through the headset as audio tones.

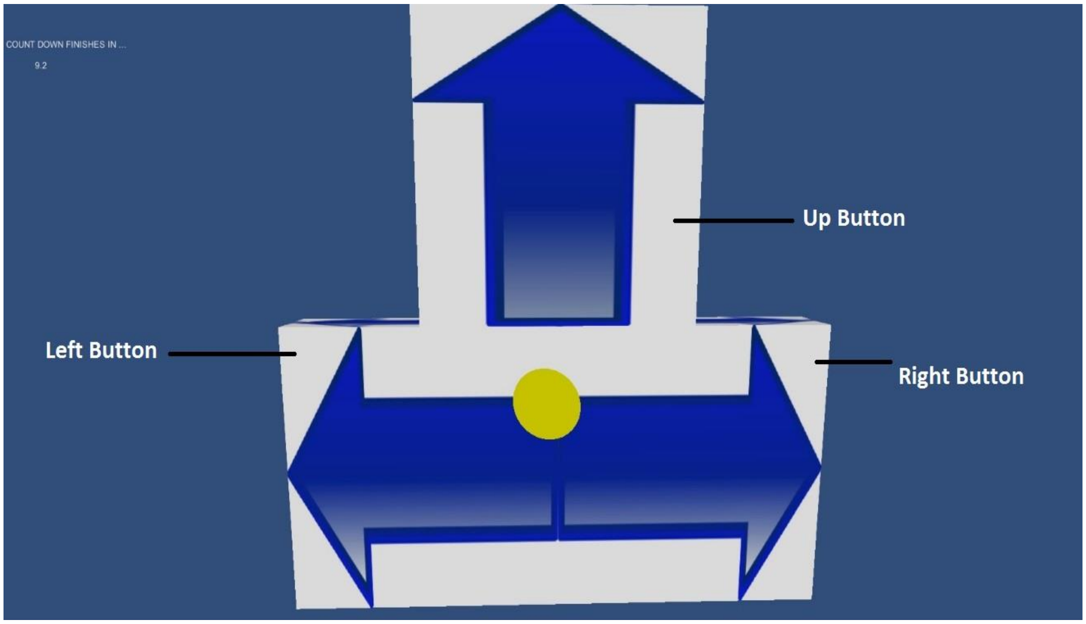


Figure 3.3: The graphical user interface (GUI) of the experiment.

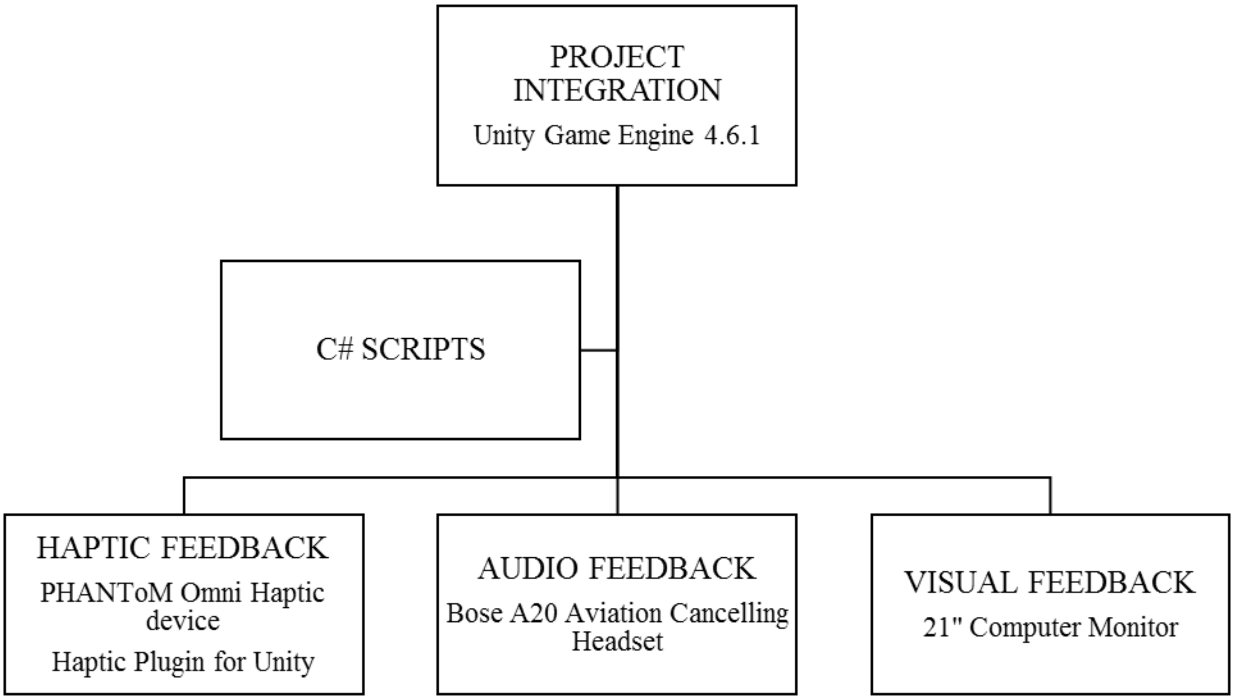


Figure 3.4: Project integration using Unity3d Game Engine.

3.3.3 Sample Size

Results from a pilot study consisting of two participants, shows a slight difference in response time between the Audio-Visual (AV) condition, Haptic-Visual (HV) and Haptic-Audio (HV) condition of the primary task as shown in Figure 3.5. The minimum response time for both participants for all conditions is 0.066 seconds.

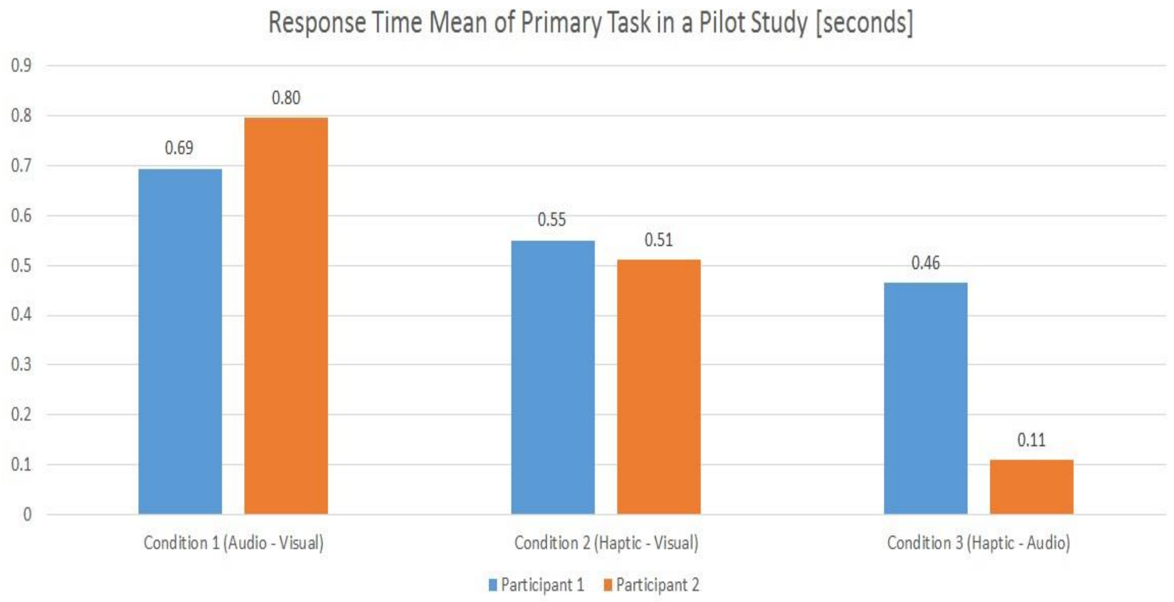


Figure 3.5: Response time mean of primary task of two participants from a pilot study.

The pilot study data was used to estimate the number of participants needed for this experiment. An iterative z-test and t-test approach was used [56]. An initial z-test was computed to approximate the degrees of freedom of the t-test using Equation 3.1.

$$n = \frac{z^2 d^2}{d^2} \qquad (3.1)$$

The confidence level of all iterations is 0.95 and is looked up in a t-statistics table for each iteration. S is the standard deviation of both participants and all conditions. The statistical power is the likelihood of the test to detect the smallest detectable value (d) and is looked up in a t-statistics table for each iteration. Multiple t-test were computed to refine the degrees of freedoms at each iteration for the smallest possible detectable value for this experiment. For the first iteration, the degrees of freedom from the z-test (Equation 2.1) is used to look up the value of the level of confidence and the value of the statistical power. For the second iteration, the degrees of freedom of the first iteration is used to look up the value of the level of confidence and statistical power. The smallest detectable value is the minimum response time from the pilot study.

The relationship is given by:

$$n = \frac{\left(t_{\alpha} - t_{\beta}\right)^2 s^2}{d^2} \tag{3.2}$$

where:

t_{α} : *Level of Confidence*

t_{β} : *Statistical Power*

s^2 : *Variance of the Difference*

d : *The Smallest Detectable Value*

df : *Degrees of Freedom of the t – test*

	Initial Z-Distribution	Iteration 1	Iteration 2
t_α	2.57	2.787	2.763
t_β	2.33	2.485	2.467
s^2	0.044	0.044	0.044
d	0.066 sec	0.066 sec	0.066 sec
df	-	25	28
Number of Participants	25	29	28

Table A.1: Sample size estimation from a pilot study.

3.3.4 Design

A between-subject design is utilized in this experiment. There are three trials; the first trial consists of a primary task only, the second trial adds a secondary cognitive, and the third trial is conducted under the same conditions as the second trial. The independent variables are the indications sent to the participant. The indications differ according to the trial in which the participant is volunteering. There are three different trials; Audio-Visual (AV), Haptic-Visual (HV) and Audio-Haptic (AH). The dependent variables are the time it takes the participants to press the virtual button (response time measured) and the number of times the participant presses the correct button (accuracy).

3.3.5 Procedure

Each participant is asked to read an information letter describing the study procedure and sign a consent form. The participant is seated and the seat is adjusted, so that their hands can reach the haptic device end-effector. The participant wears the headset and grasps the end-effector of the haptic device. The implementation of this experiment includes two different tasks: a primary task and a secondary task. During each experiment, each participant has to perform three trials.

The primary task includes three different feedback modalities; they are auditory, visual and haptic. In the primary task, two of the modalities are combined to form the indications to press the virtual buttons; the combined modalities are Audio-Visual (AV), Haptic-Visual (HV) and Audio-Haptic (AH). During each trial, every participant is introduced to only one of the combinations. Each experiment is divided into three trials. The first trial includes only the primary task while the second and third trials include the primary and secondary task.

When the first trial is started, the participant has a one-minute training to get familiar with the operation of the haptic device. The participant also has a one-minute training in the second trial. For each trial, each participant presses the virtual button that is indicated by the feedback combination. Depending on the combined feedback modes, the participant receives a flash of light from the button itself, a tone from the headset or a force from the haptic device.

The direction of the light, tone or force indicates the button to press. For instance, in the AV condition, if a participant receives a flash of light from the up button and a tone from both earbuds, the participant is required to press the up button. Figures 3.6, Figure 3.7 and Figure 3.8 shows the all possible visual, auditory and haptic feedback modes for this experiment. For each trial, a button is chosen randomly at 3 second interval. Participants are required to press all indicated buttons in the first trial. In the second and third trials, participants are required to press the buttons while performing a secondary task which is described below. All trials are 5 minutes long with 1 minute of training prior to the first and second trial.

The secondary task includes three different feedback modalities; auditory, visual and haptic. In the secondary task, one of the modalities is used to form the information channel. The only trials that include the secondary task are the second trial and third trial.

After the first trial is completed, the participant has a one-minute training to get familiar with the operation of the haptic device and the procedure of recognizing and writing the codes. For each trial, each participant performs the primary task (pressing a virtual button) while a Morse code (e.g. •— •• — — •) is conveyed to the participant through another remaining sensory feedback. Depending on the feedback mode, the participant receives flashes of light from the button itself, tones from the headset or vibrations from the haptic device. For example, a participant receives flashes of light from the virtual button when the visual feedback is chosen. All codes are between two to six bits of size and the structure is shown in Table 4.1. The encoded Morse codes used in the secondary task are presented in Table 4.2. For each trial, a code is chosen randomly every 60 seconds. Participants are required to recognize and write the code by pausing the experiment (pressing the space bar). The experiment is automatically paused every 60 seconds to give the participant a chance to write the code on the paper provided to the participant.

Distractor	Haptic	Audio	Visual
Type	Vibrations	Tones	A light flash
Actuator	Geomagic PHANToM Omni	Bose A20 headset	Computer Monitor
Duration	Short Vibrations 0.3s Represented by • Long vibrations 0.9s Represented by —	Short tones 0.3s Represented by • Long tones 0.9s Represented by —	Short light flashes 0.3s Represented by • Long light flashes 0.9s Represented by —

Table A.1: The structure of codes in the secondary task.

Code Size [bits]	**Trial 2**	**Trial 3**
2	— •	• •
3	• • —	• — —
4	• — • •	• — — •
5	— • • — —	— — — — •
6	• — — • • •	— — • — • —

Table A.2: Encoded Morse codes for Trial 2 and Trial 3.

The quantitative data collected in this experiment are the primary task response time (the time it takes a participant to press the appropriate button), response time accuracy (the number of correct button presses), the secondary task response time (the time it takes a participant to recognize the correct code), the secondary task accuracy (the number of correct codes), the press and release time (the time it takes a participant to press and release the correct button and the task completion time (the time it takes a participant to complete a trial. In this thesis, only the primary task and the secondary task response time and accuracy are analyzed.

After each trial, participants answer a questionnaire. The questionnaire evaluates the subjective performance and subjective workload for participants operating a HMI. NASA Task Load Index (TLX), as shown in Appendix C, is a multi-dimensional questionnaire designed by the Ames Research Center [57]. [57]. Participants answer questions that assess different criteria (C_i) of cognitive workload:

C1 Mental Demand (20 points): how mentally demanding was the task?

C2 Physical Demand (20 points): how physically demanding was the task?

C3 Temporal Demand (20 points): how hurried or rushed was the pace of the task?

C4 Performance (20 points): how successful were you in accomplishing what you were asked to do?

C5 Effort (20 points): how hard did you have work to accomplish what you were asked to do?

C6 Frustration (20 points): how insecure, discouraged, irritated, stressed, and annoyed were you?

Each criterion is a 20-point scale from left to the right except performance criterion. Performance has 20-point scale starting from right to left. Participants are made aware of this exception before the start of the experiment. NASA TLX has a paper-and-pencil and online version. In this experiment, the paper-and-pencil version is used.

Furthermore, NASA TLX has a weighing scheme where participates compare criteria pairwise in terms of importance. In this study, the weighing scheme is not used. However, a RAW TLX is used instead. The RAW NASA TLX sums all subscales to get a workload scale of 100-points. Many research are in favor of RAW NASA TLX since it eliminates individual biases [58]. The subjective performance is elicited from the performance criterion in the questionnaire.

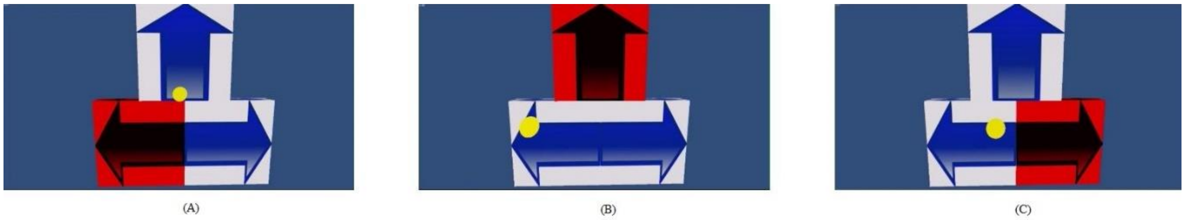


Figure 3.6: Visual feedback mode (A) Left virtual button flashing (B) Up Virtual button flashing (C) Right button virtual button flashing.

Figure 3.7: Auditory feedback mode (A) A tone from the left earbud (B) A tone from the both earbuds (C) A tone from the right earbud.

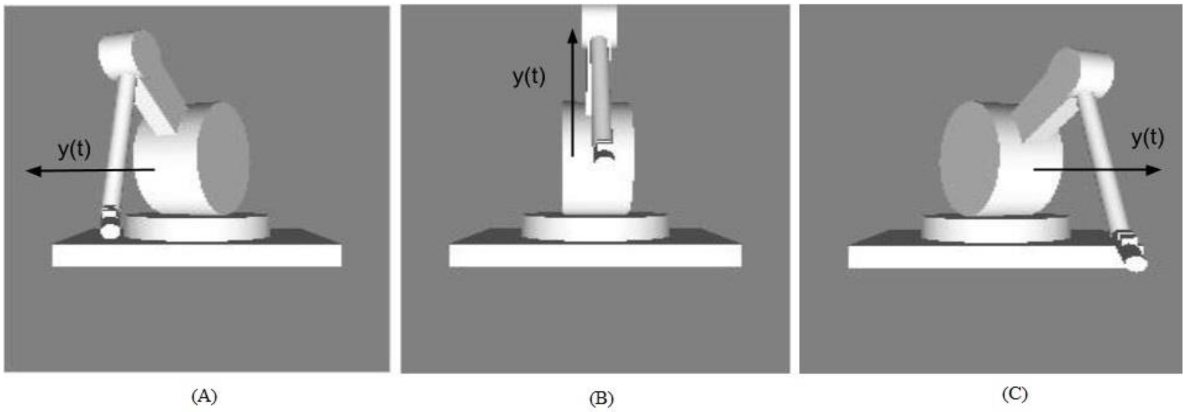


Figure 3.8: Haptic feedback mode (A) A force to the left (B) A force up (C) A force to the right.

3.4 Results

The results for the primary task and the secondary task are discussed in this section. The primary task response time for each subject is presented in Table A.1. The average primary task accuracy for each subject is presented in Table A.2. The average subjective performance is presented in Table A.18. Figure 3.9 and Figure 3.10 depict the box plot for the participants' response time, accuracy. The box plot shows the data distribution by representing the median, first and third quartiles, minimum and maximum and outliners [59]. The average secondary task response time for each subject is presented in Table A.3.The average secondary task accuracy for each participant is presented in Table A.4. The average subjective workload is presented in Table A.17. Figure 3.11 and Figure 3.12 depict the box plot for the secondary task response time and secondary task accuracy. A more detailed description of each of these results will be presented in this section.

It is important to show a specific set of data comes from a normally distributed population while performing hypothesis testing. ANOVA assume normally distributed data and is especially sensitive to non-normal data sets. Therefore, test for normality is conducted for all data sets (all trials and conditions) for this experiment. A Lilliefors test [60] of normality is computed with a Monte Carlo approximation of 1×10^4. The p-value of the Lilliefors test are included in a table of values for a significance level between 0.001 and 0.5. For more accurate values, Monte Carlo approximation is used to minimize the error due to the simulation of the p-value. The results of the tests are shown in Table B.17, Table B.18 and Table B.19 and show a normal or approximately normal distribution.

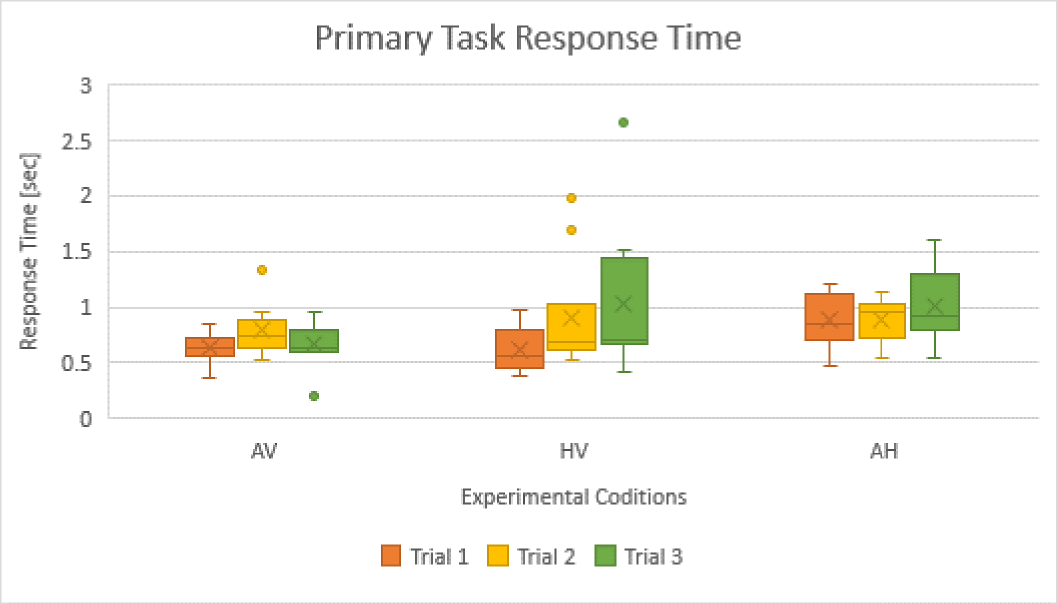


Figure 3.9: Box plot of primary task response time for all participants.

The average response time of the primary task is calculated for all trials and conditions and is shown in Table A.1. Figure 3.9 depicts the average response time of the primary task for all trials and conditions. The data point depicted as • are outlying data. Initially, the median value of response time for the AV condition, HV condition and AH condition are 0.63 seconds, 0.55 seconds and 0.85 seconds respectively. In the Trial 2, there is an increase in response time yet reserving the same trend. The HV condition has the lowest response time average and the AH condition has the highest response time average. The AV condition has a higher response time average than the HV condition and lower response time average than the AH condition. In Trial 3, the AV condition`s response time significantly decreased from 0.74 in Trial 2 seconds to 0.64 seconds in Trial 3.  The HV condition`s response time slightly increased from 0.69 seconds in Trial 2 to 0.70 in Trial 3. There is also a slight decrease in response time in the AH condition from 0.95 seconds in Trial 2 to 0.92 seconds in Trial 3.

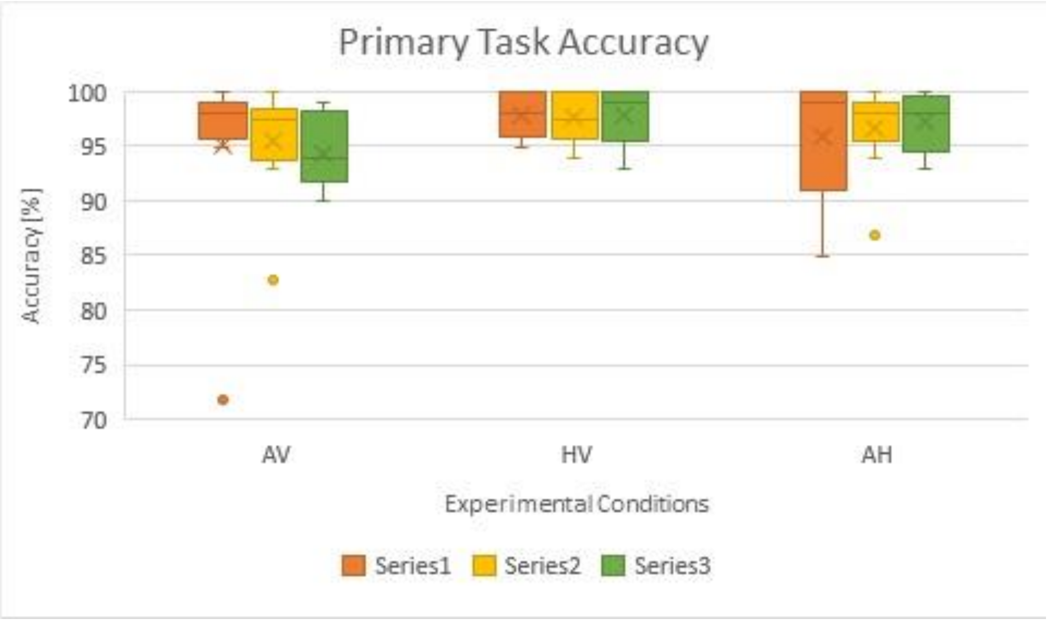


Figure 3.10: Box plot of primary task accuracy for all participants.

The accuracy is calculated for all trials and conditions. The results for all participants are shown in Table A.2. The accuracy is the number of correct button presses over the total number of signaled buttons. Figure 3.10 shows the box plots for the accuracy for each trial. The accuracy for all conditions in Trial 1 is considerably high. The accuracy for the AV and AH are 97 % and 98 % with a negligible increase of 1 %. The accuracy in Trial 2 is consistent for all conditions. 97 % accuracy is achieved for all conditions in Trial 2. In Trial 3, only the AV condition decreased to 93 %.

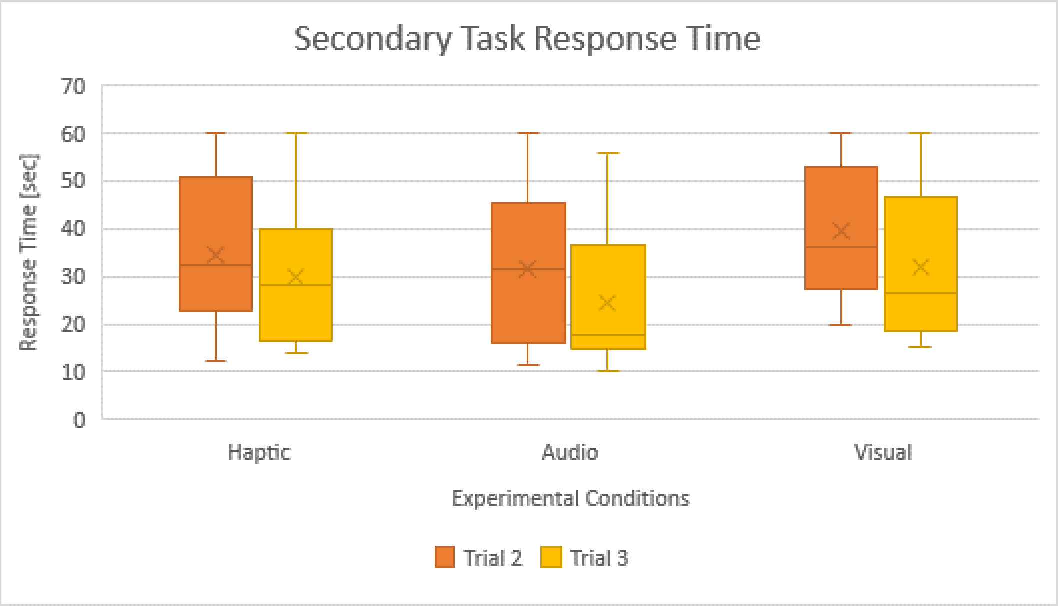


Figure 3.11: Box plot of the secondary task response time mean for Trial 2 and Trial 3.

Response time of secondary task for Trial 2 and Trial 3 is plotted in Figure 3.11. In both trials, auditory feedback is generally the lowest in terms of response time. As shown by the line of response time median, participants exposed to visual feedback in the secondary task take more time to recognize the codes. In Trial 3, however, visual feedback response time is lower than haptic feedback, yet higher than auditory feedback mode.

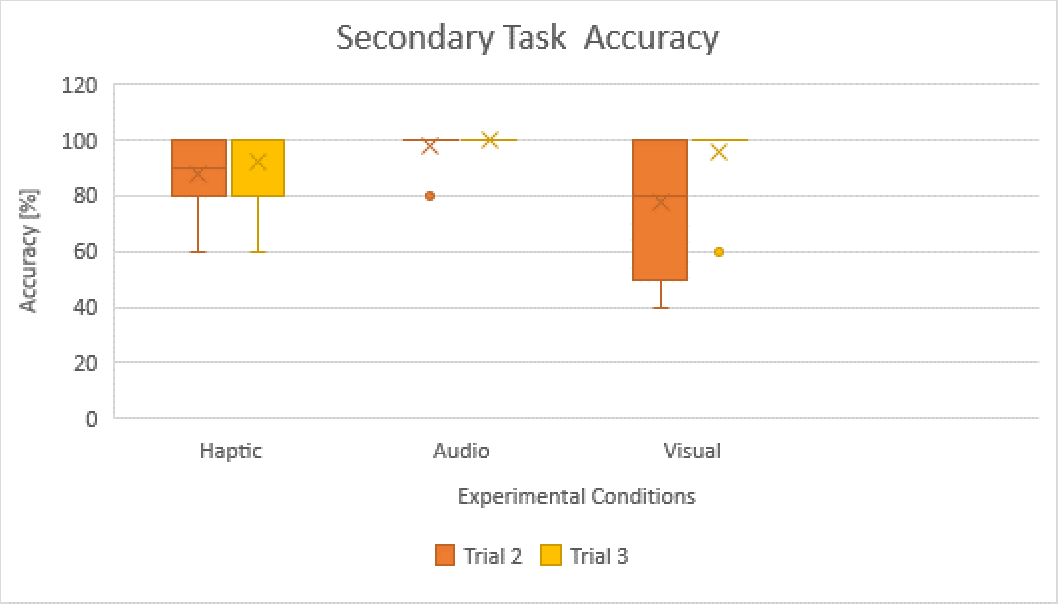


Figure 3.12: Box plot of the secondary task accuracy for Trial 2 and Trial 3.

Secondary task accuracy follows a similar pattern to the response time as plotted in Figure 3.12. The auditory feedback mode has the highest accuracy of all conditions in Trial 2 and Trial 3. The visual feedback mode has a lower accuracy than the haptic and auditory feedback modes. Even though the median accuracy of all conditions is 100 % in Trial 2, Figure 3.12 shows extreme values for the haptic condition.

The subjective performance and subjective workload are collected from the NASA TLX and converted to numerical values. The subjective workload has a 100-point scale. The subjective performance has a 20-point scale. They are given by the relationship:

$$Subjective\,Performance = C_4 \tag{3.3}$$

$$Subjective\,Workload = \sum_{i=1}^{6} C_i \tag{3.4}$$

The subjective performance for all participants is elicited from the NASA TLX questionnaire and is shown in Table A.17. The subjective performance has a 20-point scale. Figure 3.13 shows the box plots for subjective performance for all trials. The AH condition has the lowest perceived performance for all trials. The HV condition has the highest perceived performance for all trials. It can also be seen that the AV condition has a significant decrease in perceived performance.

The median subjective workload for all participants is shown in Figure 3.14. The maximum subjective workload reported is in Trial 2. The subjective workload for the AV condition is the highest in Trial 1 and Trial 3. Because of the introduction of the secondary task, the subjective workload of the AH condition increased significantly. The HV condition has the lowest subjective workload between conditions in all trials.

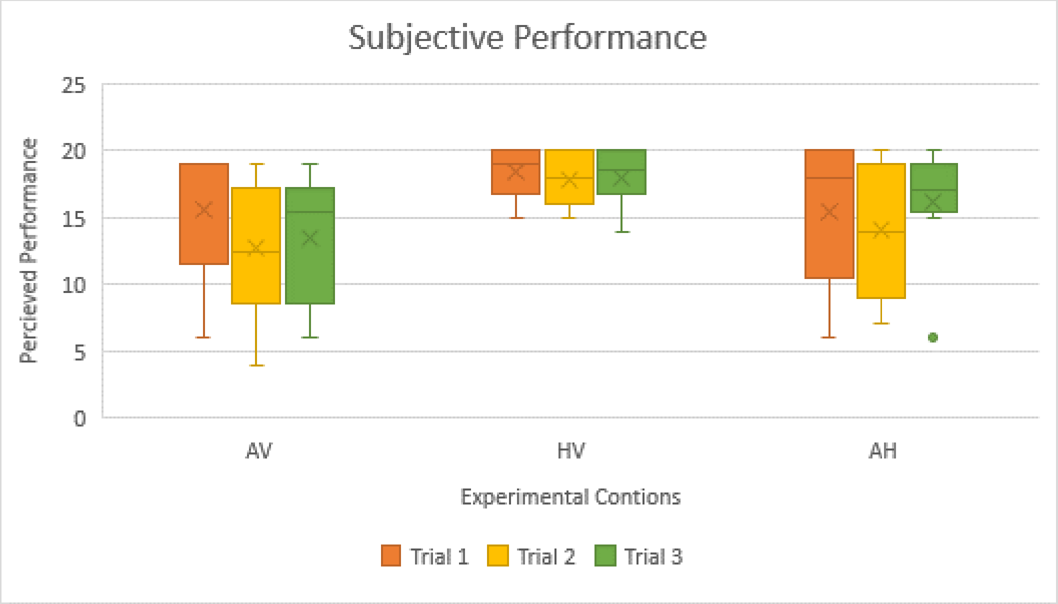


Figure 3.13: Box plot of participant`s subjective performance for all participants.

Figure 3.14: Box plot of participant`s subjective performance for all participants.

3.5 Discussion

The results of the primary task response time, primary task accuracy, secondary task response time, secondary task accuracy, subjective performance and subjective workload are further analyzed. Grubb's test [61] for detecting outlying data points is computed to confirm the data points which are outliers in Section 3.4. Grubb's test takes a data point from a normally distributed data set. Since the method tests one data point at a time, an iterative method has been implemented to detect outlying data points. The critical value of the Grubb's test is compared to the two-sided hypothesis test of Equation 3.6. If the critical value of the Grubb's test is higher than the two-sided test, the hypothesis of that the data point is not outlier is rejected and the data point is removed. All the data points in Figure 3.9, Figure 3.10, Figure 3.12 and Figure 3.13 that are denoted by [•] are removed using Grubb's test. The relationship of the Grubb's test is given by:

$$T_n = \frac{x_n - \tilde{x}}{s} \tag{3.5}$$

$$T_n > \frac{N-1}{\sqrt{N}} \sqrt{\frac{\left(t_{\alpha/(2N),N-2}\right)^2}{N-2+\left(t_{\alpha/(2N),N-2}\right)^2}} \tag{3.6}$$

where:

T_n : *The critical value of the Grubb's test.*

x_n : *The Value of the nth Observation.*

$\tilde{x}$: *Arthimetic Average of the Data Set.*

s : *The* Standard Deviatoin *of the Data Set.*

$t_{\alpha/(2N),N-2}$: *The Upper Critical Value of the t – Distrubution.*

$\alpha/(2N)$: *Significance Level.*

N : *Sample Size (Size of the Data Set).*

The results are analyzed using a between-subject Analysis of Variance (ANOVA) and multiple two-sample t-tests. The analysis is done at 0.05 significance level. In some cases, the analysis is done at 0.10 or 0.15 significance level. The significance level is stated when it is higher than 0.05. The average response time, accuracy, and subjective performance is calculated for all trials and conditions after eliminating all outlying data points. Table 3.2 summarizes the average primary task response time and standard deviation for all trials.

Trial	AV Condition		HV Condition		AH Condition	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
1	0.6366	0.0733	0.6126	0.1948	0.8776	0.2400
2	0.7430	0.1307	0.6673	0.0895	0.8882	0.1926
3	0.7180	0.1321	0.8438	0.3690	1.0121	0.3433

Table 3.2: The average response time (sec) and standard deviation of all trials.

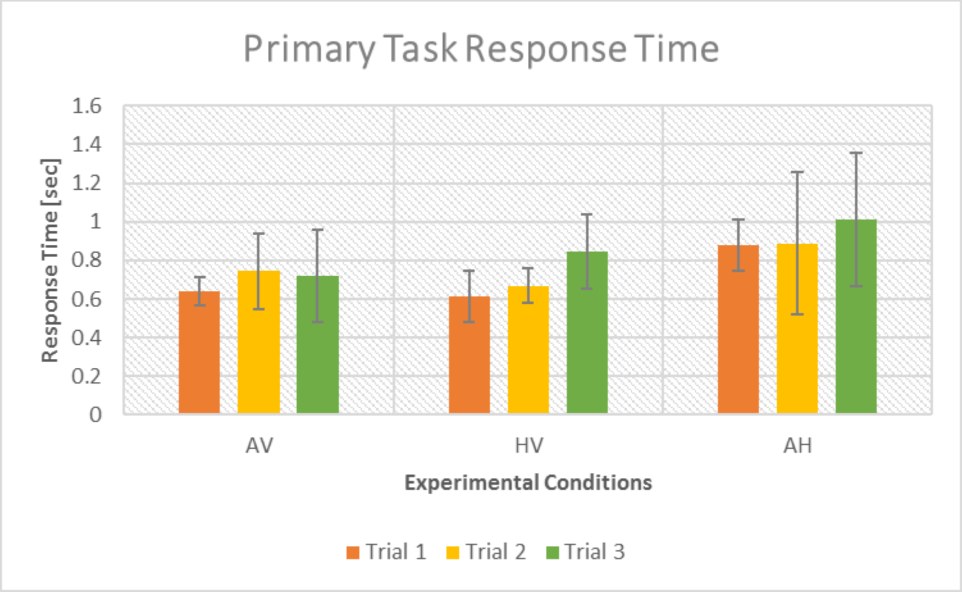


Figure 3.15: The average primary task response time (sec) for all participants.

Figure 3.15 shows the average response time after deleting outliners for all trials. Figure 3.15 shows a similar trend to the data in Figure 3.9. The AH condition shows the highest response time for all trials. The HV condition has a lower response time than the AV condition in Trial 1 and Trial 2. However, the average response time for the AH condition significantly increases in Trial 3.

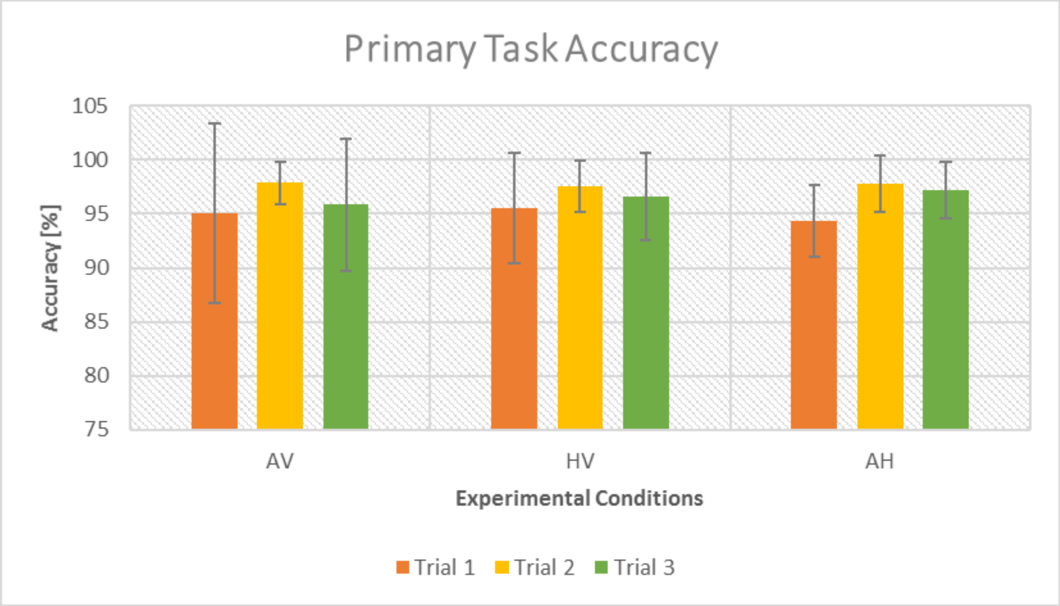


Figure 3.16: The average primary task accuracy (%) for all participants.

The accuracy for all trials is shown in Figure 3.16. Accuracy for all trials is almost perfect for all conditions. There is a slight decrease in accuracy for the AH condition in Trial 1. There is also an insignificant decrease in accuracy for the AV condition in Trial 3. Subjective performance also indicates a similar trend to the unfiltered data. The highest subjective performance in the HV condition can be seen in Figure 3.19. The AV condition has a higher performance in Trial 1 while it is lower in Trial 2 and Trial 3. As shown in Figure 3.15, Figure 3.16 and Figure 3.19, there is a common tendency for some conditions to have higher response time and a lower accuracy. This is also apparent in participant's response to subjective performance.

Table B.18 and Table B.19 lists the results of the Lilliefors test of normality for the secondary task response time, secondary task accuracy and subjective workload. According to the table, data shows normal or approximately normal distribution. Moreover, one-way ANOVA and multiple two-sample tests for the secondary task are listed in Appendix B. Figure 3.17 Figure 3.18 illustrates the average secondary task response time, the average secondary task accuracy and the average subjective performance, respectively, after deleting outlying data. The results show a similar trend to the data prior to outlier elimination.

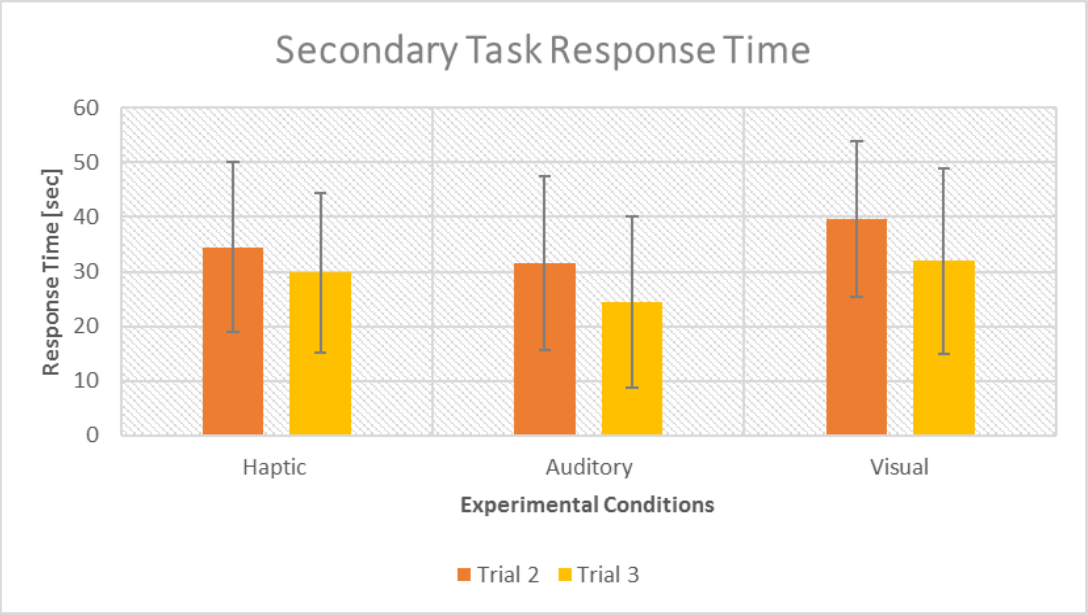


Figure 3.17: The average secondary task response time (sec) for all participants.

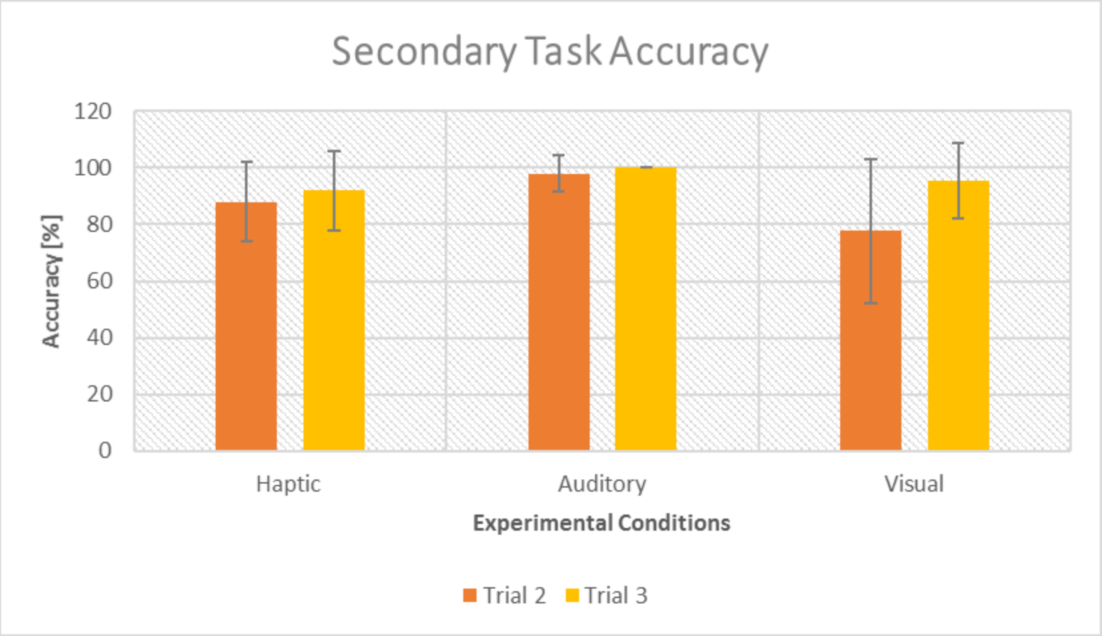


Figure 3.18: The average secondary task accuracy (%) for all participants.

Figure 3.19 shows the average subjective performance for all the trials. The average subjective performance of the HV condition is the highest for all trials. This is also true for the response time of the primary task and supports that the HV condition increases the task performance. In Figure 3.20, the subjective workload for all trials is depicted. As can be seen in Figure 3.20, the subjective workload for the HV condition is the lowest and the AH condition has the highest perceived workload for most of the trials.

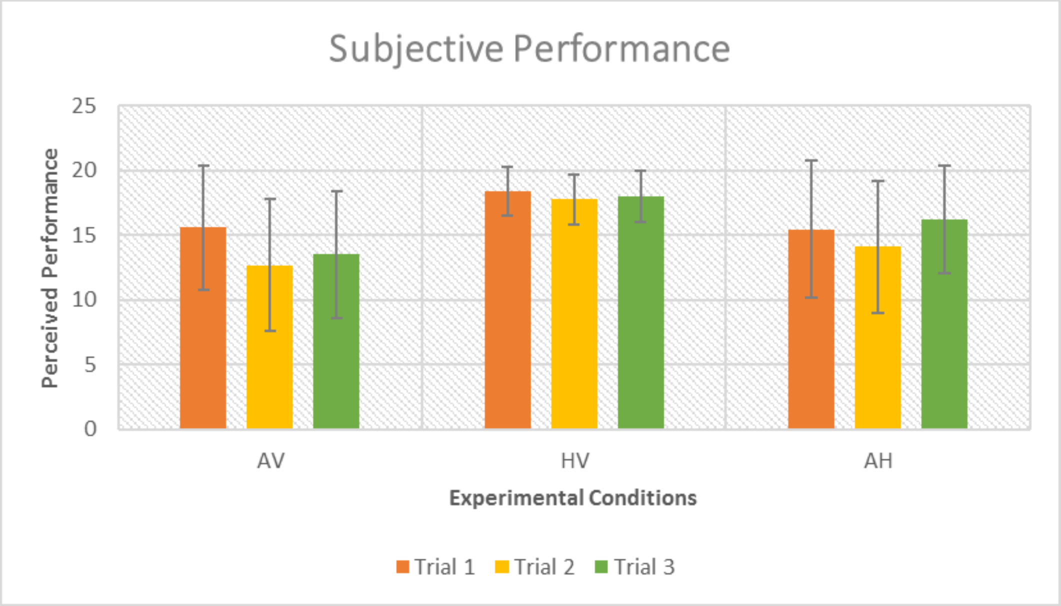


Figure 3.19: The average subjective performance for all participants.

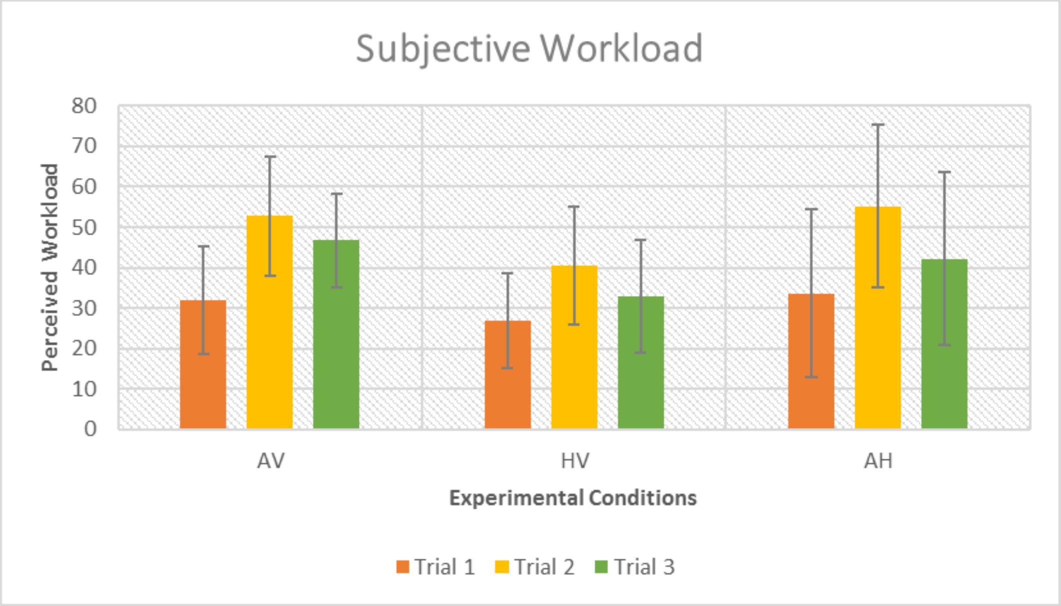


Figure 3.20: The average primary and secondary task subjective workload for all participants.

3.5.1 Tukey Honest Significant Difference (HSD) Pairwise Comparison

Figure 3.15 shows a significant difference in the primary task response time between the AH condition and the HV condition. A slight difference between the HV condition and AV condition can also be seen in Figure 3.15. The results of the ANOVA test, p = 0.0102, validate the alternate hypothesis that there is a difference between different sensory modalities. Tukey Honest Significant Difference (HSD) pairwise comparison [62] of the ANOVA test statistic is performed. Figure 3.21 shows the pairwise comparison of Trial 1. The response time for pressing a virtual button does not include any cognitive loading as there is not a secondary task for Trial 1.

The Tukey HSD multiple comparisons does not show any difference between the AV condition and the HV condition; however, it shows some differences between other conditions. It seems that participant's responses are sluggish for audio and haptic feedback mode. Number of factors might contribute to the increase of response time and decrease in accuracy for participants using audio and haptic force feedback. Visual displays are predominantly used by participants. Consequently, the absence of visual cues is the most influential factor.

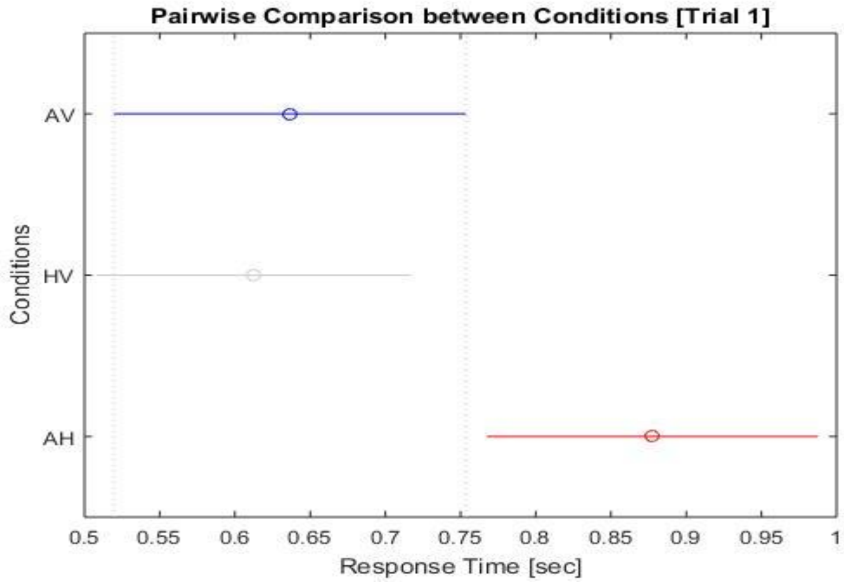


Figure 3.21: A pairwise comparison of the response time mean in Trial 1.

Because of the introduction of secondary task, there is a slight increase in response time especially apparent in AV condition and HV condition. The effects of the secondary task in the response time and accuracy of the primary task is discussed in Chapter 4. In Trial 2, the null hypothesis that all conditions are the same is also rejected at 0.05 alpha level. P-value of 0.01 confirms that there is a difference in response time between conditions. Figure 3.22 shows the Tukey HSD pairwise comparison for Trial 2. Similar to Trial 1, the major difference in response time lies between the AH condition and the HV condition. The two sample t-test adds a great confidence (p = 0.0097) to reject the null hypothesis. As can be seen, the difference between the AH condition AV condition is not signficant. However, the difference can be virtually noticed from the two sample t-test with a p-value equal to 0.07. The low p-value (close to p = 0.05) indicates that a difference can be revealed at a higher significance level. Another notable result is the minor difference between the AV and HV conditions. Although the response time increased, a difference cannot be determined since the p-value for the t-test is 0.18.

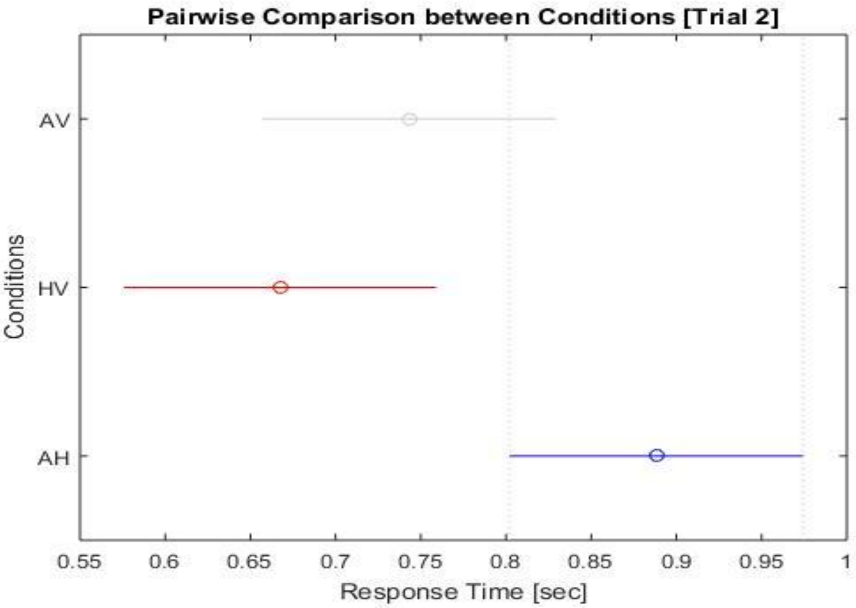


Figure 3.22: A pairwise comparison of the response time mean in Trial 2.

Figure 3.23 shows the Tukey HSD of the primary task response time for Trial 3 at a significance level of 0.15. According to the figure, there is a difference in response time between the AH and AV conditions. The ANOVA test p-value is 0.1364, and the two-sample t-test p-value for the difference between the AH and AV conditions is 0.0290. It is worthwhile to note that the response time considerably increased for the HV condition from Trial 2 to Trial 3. The Tukey HSD and the two sample t-test do not show any evidence for a difference between the AV and HV conditions (p-value of two sample t-test >> 0.05) and HV and AH conditions (p-value of two sample t-test >> 0.05).

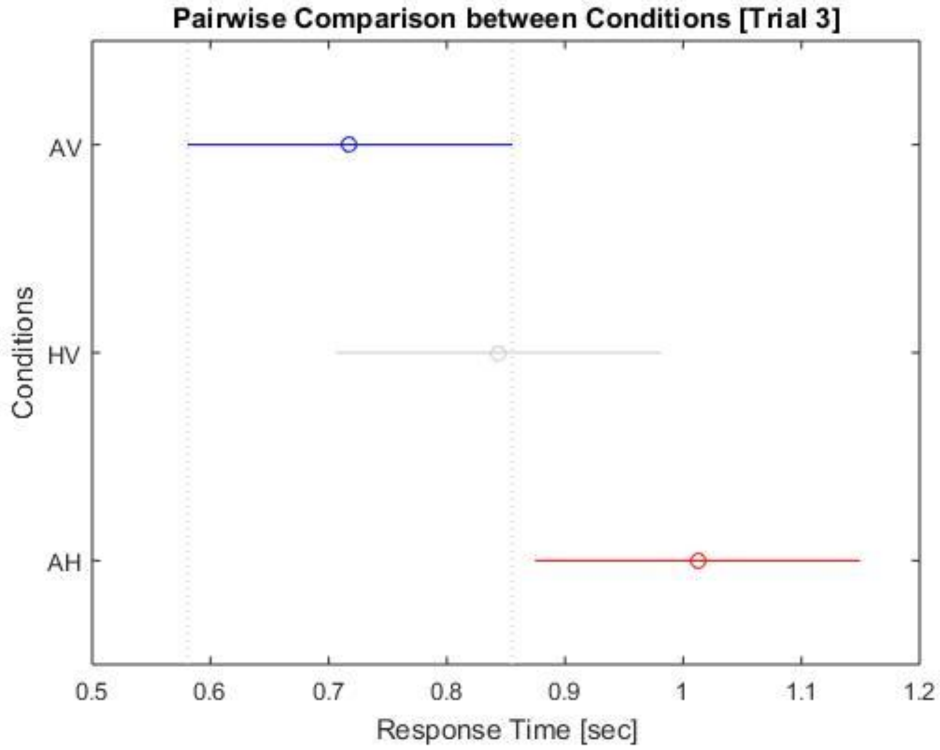


Figure 3.23: A pairwise comparison of the response time mean in Trial 3 at 0.15 confidence level.

Subjective performance responses from the NASA TLX are also analyzed using one-way ANOVA and two-sample t-tests. The results for the ANOVA and two-sample t-tests are provided in Table B.11, Table B.12, Table, B.13 and Table B.22. Trial 1 does not show any difference in perceived performance. This might be attributable to the absence of the cognitive loading on participants. According to Figure 3.24 and Figure 3.25, perceived performance of AV condition is not affected in Trial 3, and it is the same for Trial 2. The perceived performance of HV condition is the highest, and it shows a difference compared to AV condition with an ANOVA p-value of 0.03 and 0.006 for Trial 2 and Trial 3, respectively.

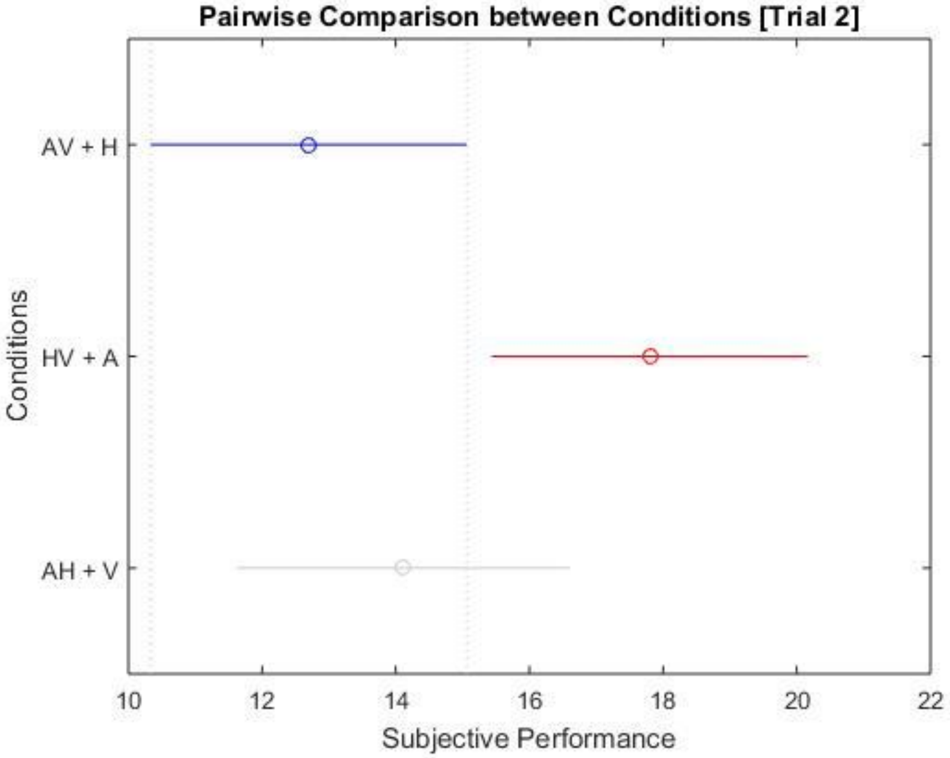


Figure 3.24: A pairwise comparison of the subjective performance in Trial 2.

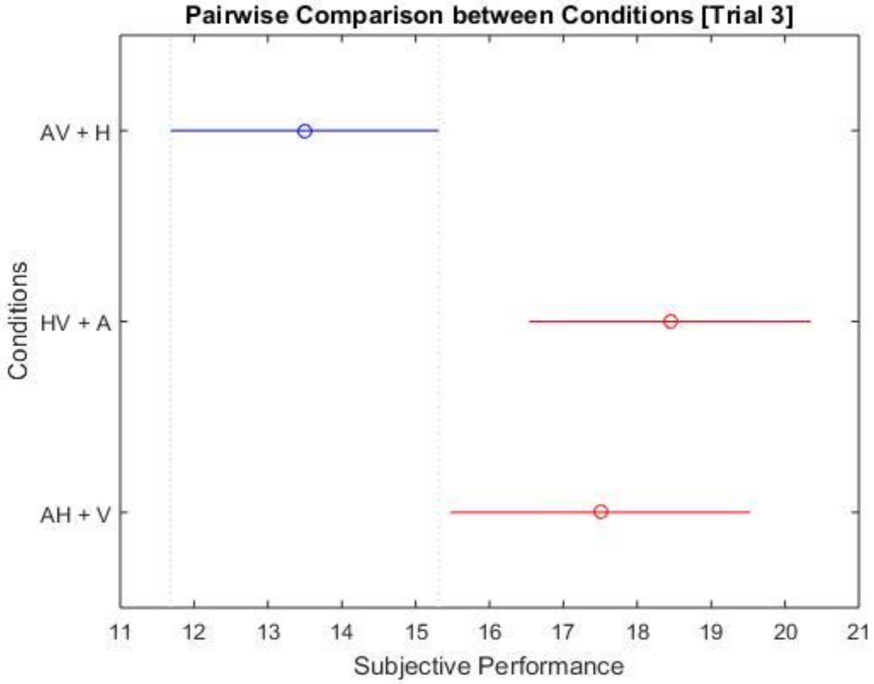


Figure 3.25: A pairwise comparison of the subjective performance in Trial 3.

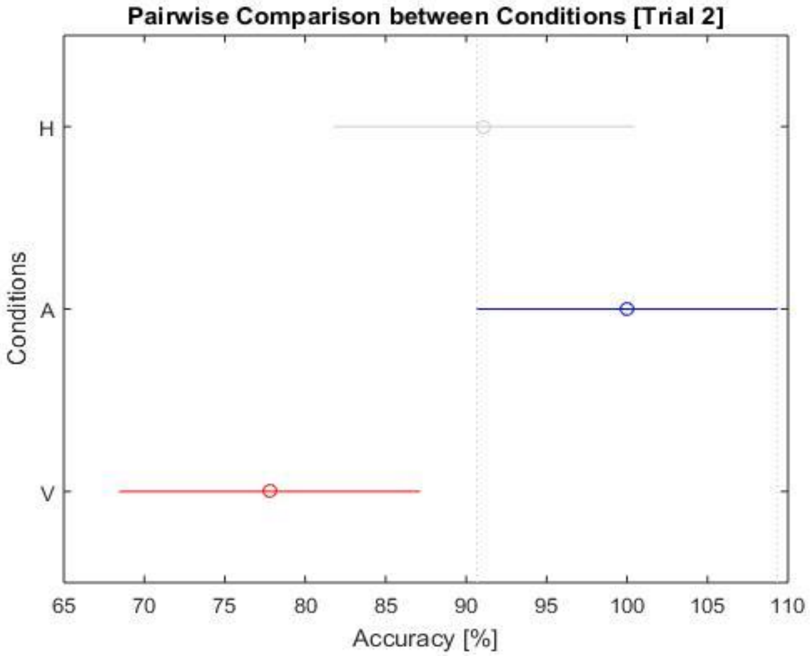


Figure 3.26: A pairwise comparison of secondary task accuracy in Trial 2.

Tukey HSD pairwise comparison method is also used in the analysis for cognitive workload to reveal any differences between feedback modalities. One-way ANOVA analysis does not show any differences in the average response time among conditions in both Trial 2 and Trial 3 (p-value >> 0.05). Figure 3.27 depicts the secondary task accuracy differences between feedback modalities in Trial 2. As can be seen from the figure, there is a difference between the visual condition and the auditory condition in terms of accuracy in identifying the codes (p-value = 0.02). The auditory condition shows a higher accuracy than the visual condition. A difference is not detected between the visual and haptic conditions or between the auditory condition and the haptic condition. In Trial 3, the ANOVA test shows the same trend at a 0.10 significance level (p = 0.07). Additionally, ANOVA analysis of subjective workload reveals differences between conditions at 0.10 significance level.

Two-sample t-test analysis of the secondary task accuracy and subjective workload supports the one-way ANOVA. The two-sample t-test shows that there is a difference between the haptic and auditory conditions (p-value = 0.02) and the visual and auditory conditions (p-value = 0.01). Moreover, there are differences in subjective workload, at a 0.10 significance level, between the AV condition and HV condition (p-value = 0.0848) and the AV condition and HV condition (p-value = 0.07). In Trial 3, the two-sample t-test shows a difference in subjective performance between the AV condition and the HV condition with a p-value of 0.02.

3.6 Conclusion

In this chapter, a study is presented to estimate the participants` performance in a virtual button-pressing task. This study shows the distinction and quality of response among three different two-feedback-modality combinations. The trend of data, which is significantly supported by the results of a one-way ANOVA, confirmed the alternate hypothesis that a certain combination of feedback (Audio-Visual, Haptic-Audio or Haptic-Visual) increases performance and reduces cognitive workload, indicating that there is at least one combination different in terms of performance (HV condition) and cognitive workload (Auditory) than other feedback combinations.

The results from the primary task indicates that null hypothesis can be rejected for the response time in Trial 1, Trial 2 and Trial 3. Additionally, the null hypothesis can be rejected for the subjective performance in Trial 2 and Trial 3. The null hypothesis for participants` accuracy cannot be rejected. As can be seen from Figure 3.16, participants perform the primary task accurately; however, the response time is affected for different conditions.

Many researchers find that only using audio and visual communication is ineffective. In [23], it is found that haptic communication increases presence and enhances user experience. Similarly, it is can be seen that haptic coupled with visual feedback has the lowest response time in the primary task. It is also evident from the subjective performance analysis that users prefer the haptic and visual communication to audio and visual and audio and haptic communication. Although haptic feedback can increase human performance, the absence of visual feedback can be problematic [9]. This trend can be seen in Figure 3.15, Figure 3.16 and Figure 3.19. The Audio-Haptic feedback has the highest response time.

Moreover, most of the ANOVA and t-tests for the analysis of cognitive workload are in favor of the null hypothesis. However, in some cases, the null hypothesis can confidently be rejected. For instance, the null hypothesis is rejected for the secondary task accuracy in Trial 2. The null hypothesis is also rejected for the subjective workload in Trial 2 and Trial 3. Unlike human performance analysis, cognitive workload does not show any difference.

Cognitive workload is commonplace for users involved in a demanding task [63]. In this section, the perceived workload of sensory modalities is tested. Overall, visual feedback poses higher workload than any other sensory modality. This is apparent from the secondary task response time, secondary task accuracy and the subjective workload.

Recent research supports the use of haptic as a communication medium [63]–[65]. Instead of overloading the visual and auditory communication mediums, haptic cues are conveyed to users fixated in a demanding task. For instance, a haptic turn-taking protocol is suggested by [65].

Additionally, utilizing visual feedback poses more cognitive workload as graphical user interface become more complex [66]. It is found that increasing the dimensionality from 2D to 3D in visual interfaces decreases the ability for participants to locate, interact and manipulate objects. To address this problem, the implementation of another feedback mode is necessary to reduce the amount of cognitive workload and ensure effective communication.

Chapter 4

The Effects of Loading on the Primary Task

4.1 Introduction

The user's performance in accomplishing a task is dependent on the type of sensory feedback the user is receiving. It has been confirmed by ANOVA test that there are differences between the means of different combined sensory modalities for the primary task. Nonetheless, the effect of cognitive workload imposed by the secondary task is not clear. Thus, in this chapter, two-way ANOVA is computed between Trial 1 and Trial 2, where the secondary task is integrated. In addition, the means of the pooled response times of the primary task from Trial 2 and Trial 3 are plotted and analyzed. The individual codes' primary task response times and accuracy, and the individual codes' secondary task response time and accuracy are also analyzed using a one-way ANOVA.

In a two-way ANOVA, two independent variables are analyzed each involving multiple levels. The continuous dependant variable is the response time mean of the primary task. The independent variables are the conditions of the primary task in Trial 1 and the conditions on the primary task of Trial 2. The conditions for the primary task of

Trial 2 are independent from the conditions for the primary task of Trial 1 because the secondary task is integrated into the Trial 2. The two-way ANOVA has three types of effects. The effects of independent variable which are the conditions of the primary task (Factor 1), the effects of the second independent variable (Factor 2: with/without the effect of the secondary task) and the interaction between the two independent variables (Factor 1 and Factor 2).

The purpose of this chapter is to study the effects of the secondary task on the variability of the response time among different levels (experimental conditions; AV, HA or HV). This phenomenon is called interaction. Since a one-way ANOVA does not offer this feature, a two-way ANOVA is conducted. This chapter also aims at determining the effects of different levels of Morse codes on primary task response time, primary task accuracy, the secondary task response time and the secondary task accuracy. Consequently, the response time and accuracy of the primary and the secondary task from Trial 2 and Trial 3 for individual codes are isolated and analyzed separately.

4.2 Hypothesis

The hypotheses of the two-way ANOVA are stated as follows:

- $H3a_{null}$: There is NOT any significant effect of Factor 1 on the variation of the dependent variable.
- $H3a_{alternative}$: There is a significant effect of Factor 1 on the variation of the dependent variable.
- $H3b_{null}$: There is NOT any significant effect of Factor 2 on the variation of the dependent variable.
- $H3b_{alternative}$: There is a significant effect of Factor 2 on the variation of the dependent variable.
- $H3c_{null}$: There is NOT any interaction between Factor 1 and Factor 2.
- $H3c_{alternative}$: There is an interaction between Factor 1 and Factor 2.

4.3 Primary Task and Secondary Task Interaction

Figure 4.1 shows the primary task average response time. Figure 4.1 indicates that both AV’s and HV’s response times increase with added secondary task in Trial 2. In Trial 3, the AV’s and HV’s response times radically spike upward. They also provide an approximately equal response times in the Trial 2 and Trial 3. The response time of the AH condition, however, increases in Trial 2 and decreases in Trial 3 showing a learning process.

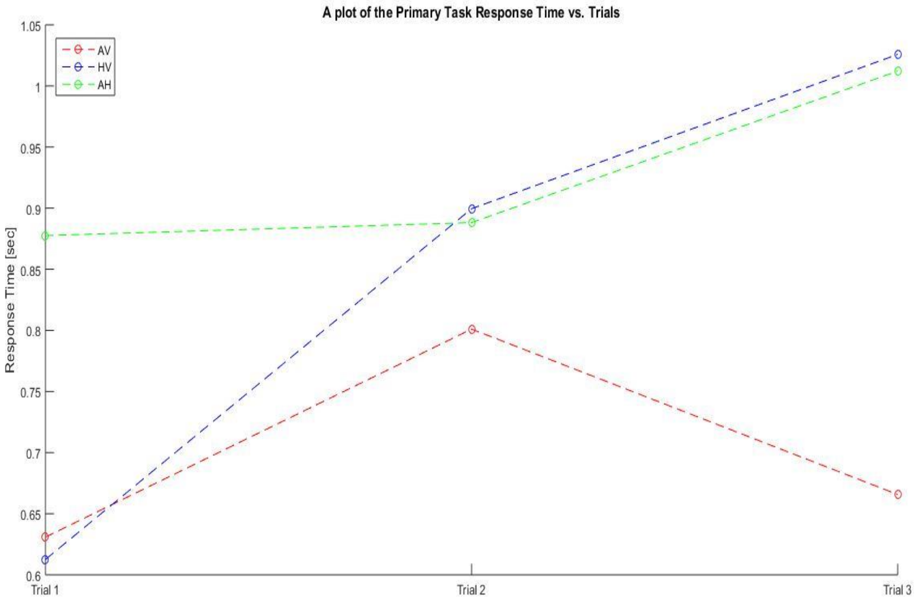


Figure 4.1: The primary task response time for the three trials

It is worthwhile to note, from Figure 4.1, the effects of the added cognitive load on the efficiency of the primary task. Thus, a two-way ANOVA is conducted to confirm the hypothesis that there are differences in response time means as a result of the Factor 2. Table B.23 summarizes the results of the two-way ANOVA. A p-value of 0.101 indicates that there is not a significant difference in response time due to the effects of Factor 1. In other words, the variation between conditions has little to do with the added cognitive task (secondary task). The p-value of Factor 2 is relatively small. This might indicate that an effect is present; however, this effect is very small to detect. For a higher significance level, say 0.15, the null hypothesis can be rejected. The effects of Factor 2 can be more salient if the level of cognitive workload is increased by, for example, increasing the code size of the secondary task or decreasing the time to recognize the code

With a p-value of 0.05, the two-way ANOVA confirms the hypothesis that there is a sufficient effect of Factor 1 on the variations of the dependent variable. We can conclude that (p-value is slightly more than 0.05) as shown in Figure 4.2. We can safely assume that there is not any interaction between Factor 2 and Factor 1 (p-value >> 0.05). There is not any combined effect from both factors influencing the average response time; however, condition variation has the most effect in the response time differences.

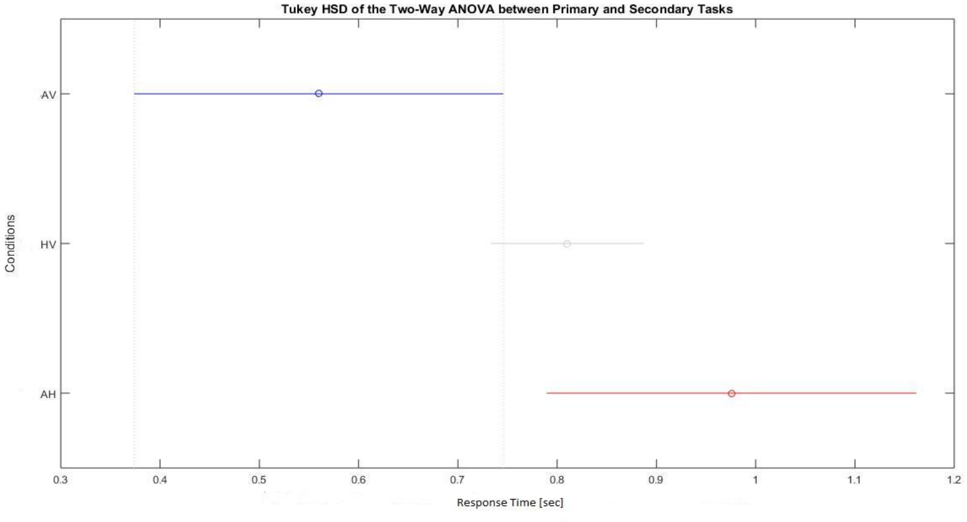


Figure 4.2: Multiple comparisons of the two-way ANOVA test statistics.

Two-Way ANOVA of Pooled Data

Figure 4.3 shows the response time mean of the primary task for pooled data from Trial 2 and Trial 3. The data is pooled to acquire more statistical power. Increasing the sample size is one way to increase the statistical power and thus increasing the probability of rejecting the null hypothesis. In this chapter, the effects of the secondary task on the variation of the dependent continuous variable are analyzed. The trends of the pooled data from Trial 2 and Trial 3 looks similar to the trends of the response time mean of Trial 1, Trial 2 and Trial 3 of the individual data sets (Figure 3.15). The figure shows that the AH condition has a higher response time for all the data from both trials. It confirms that the secondary task has a little effect on response time. A two-way ANOVA, as shown on Table B.24, has also been conducted for the pooled data to confirm whether Factor 2 has an effect on primary task response time. A p-value of 0.02 confirms that there is a significant effect of the secondary task on primary task performance; however, Factor 2 does not effect the trends of the primary task as shown in Figure 4.3. Factor 2 gives a higher response time for all conditions compared to Figure 4.2 for the two-way ANOVA between Trial 1 and Trial 2 alone.

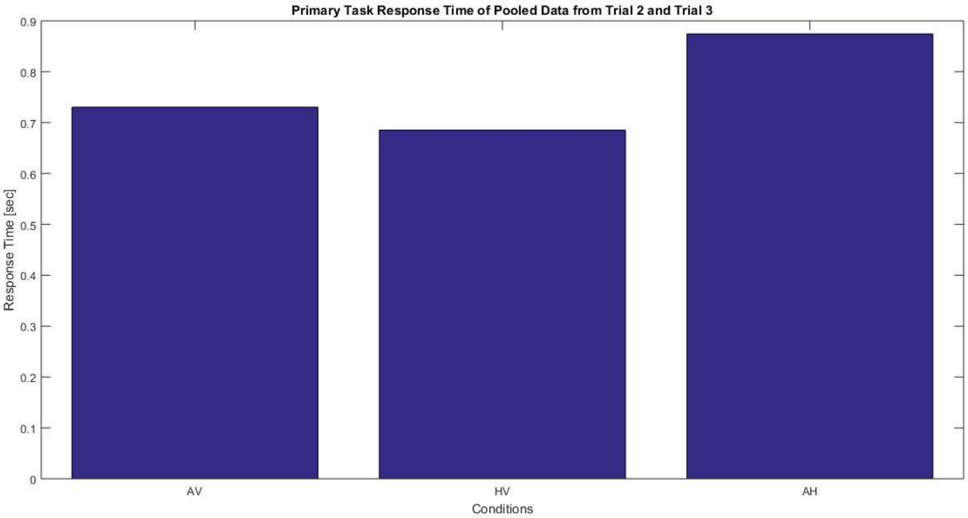


Figure 4.3: Primary task response time of pooled data from Trial 2 and Trial 3.

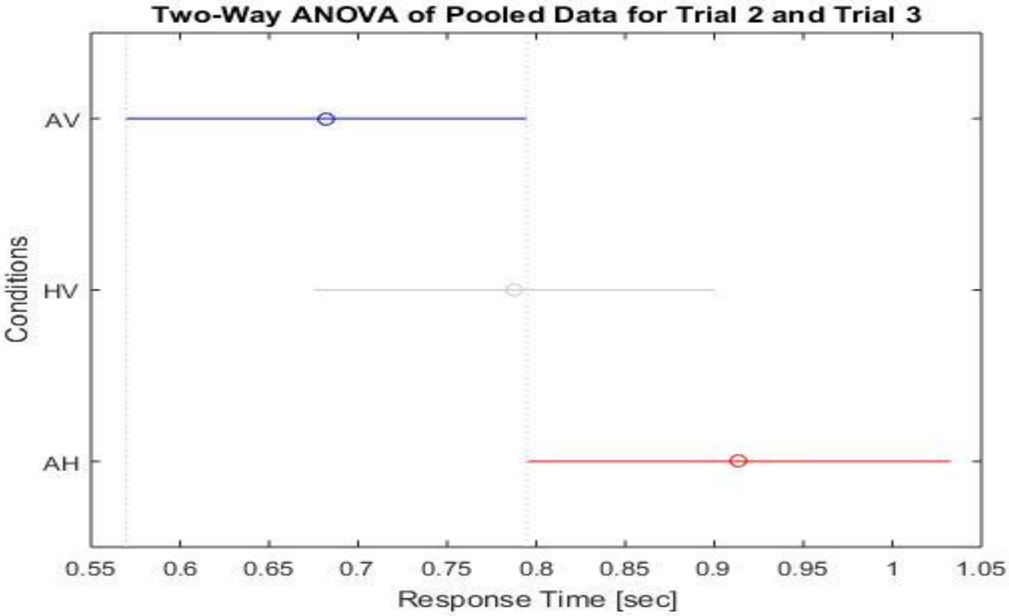


Figure 4.4: Multiple comparisons of the two-way ANOVA test statistics between Trial 1 and the pooled data of Trial 2 and Trial 3.

4.4 Analysis of Performance and Workload for Individual Codes

Individual code performance and workload are analyzed using the same method from Chapter 3 and Chapter 4. Figure 4.4 shows the response time of the primary task and the corresponding response time of the secondary task for each code. As can be seen, as the code size increases, the response time slightly increases for the AV condition (approximately 0.04 seconds) and the HV condition (approximately 0.02 seconds) with few exceptions. In the AH condition, the response time is almost constant. In the secondary task, the change in response time is drastic for the 6 bits of information; however, the response time for the 4 and 5 codes are not consistent. The tendency of the 6 bits codes to have a drastic increase in response time is in agreement with Miller's experiment [51].

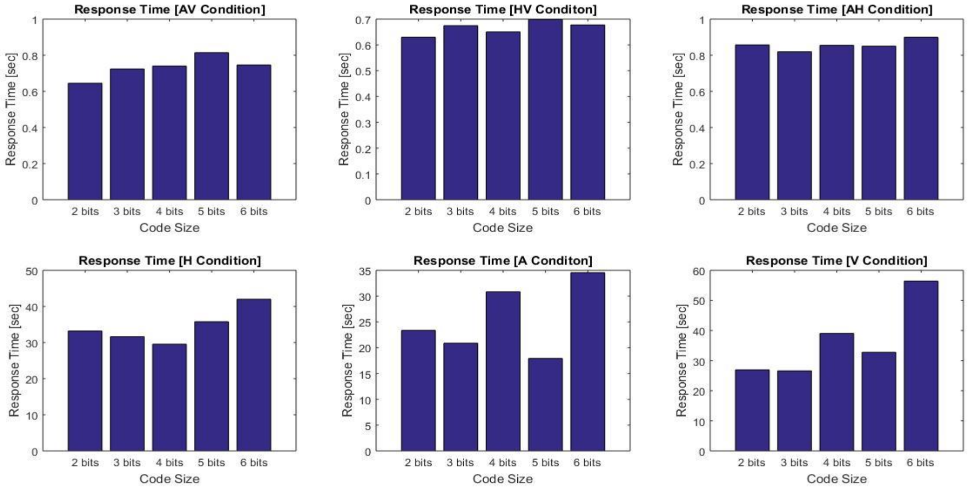


Figure 4.5: The average primary and secondary response time (sec) for individual code

Tukey HSD pairwise comparison for the response time of the secondary task (Visual condition) is depicted in Figure 4.5. It is worthwhile to note the dramatic increase in response time for the 6 bits codes. This finding supports the outcomes of Miller's experimentation in "absolute judgement" [51] that users can only hold up to 7 items in their working memory.

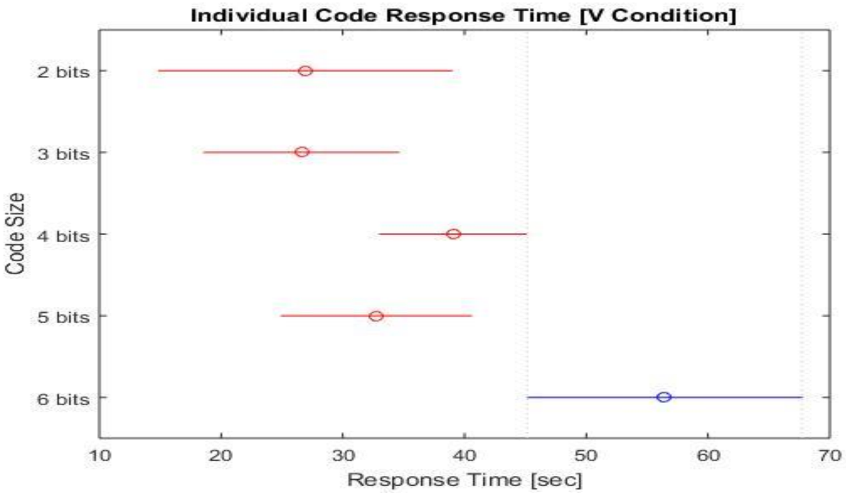


Figure 4.6: A pairwise comparison of the secondary task Visual condition response time (sec) for individual codes.

The results of the individual code accuracy are shown in Figure 4.6. The overall primary task decrease in accuracy for larger codes can be attributable to the high workload posed on participants by the secondary task. This tendency is present in the haptic and visual conditions. In Figure 4.6, primary task accuracy decreased significantly to approximately 86% in the AV condition and 93% in the HV condition when participants were exposed to 6 bits codes in the primary task. Additionally, the Tukey HSD for the secondary task accuracy shows a lower accuracy for the 6 bit codes for the haptic condition and the visual condition as shown in Figure 4.9 and Figure 4.10, respectively.

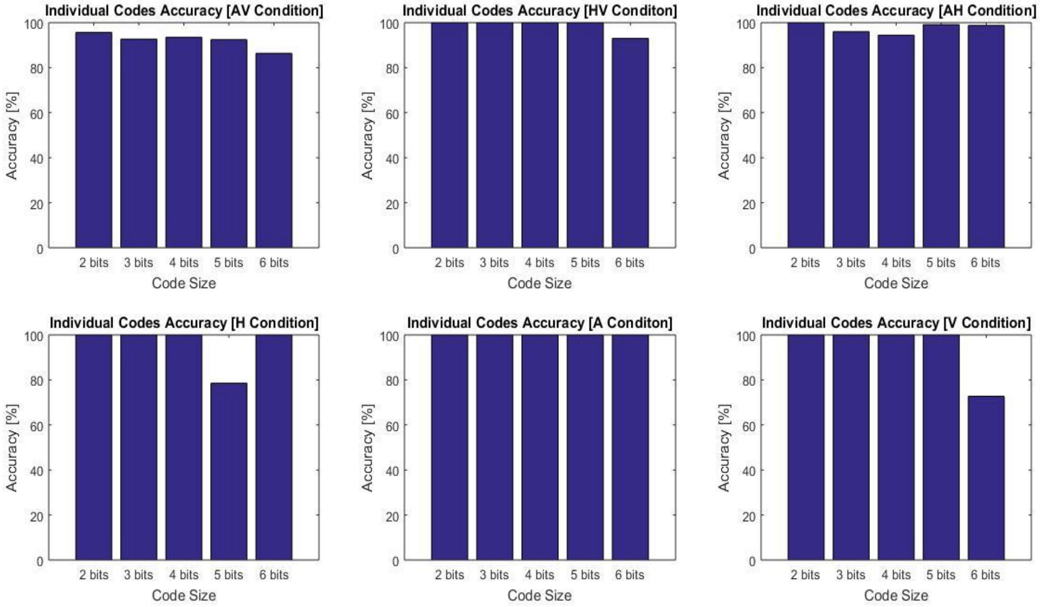


Figure 4.7: The average primary and secondary task accuracy (%) for individual codes.

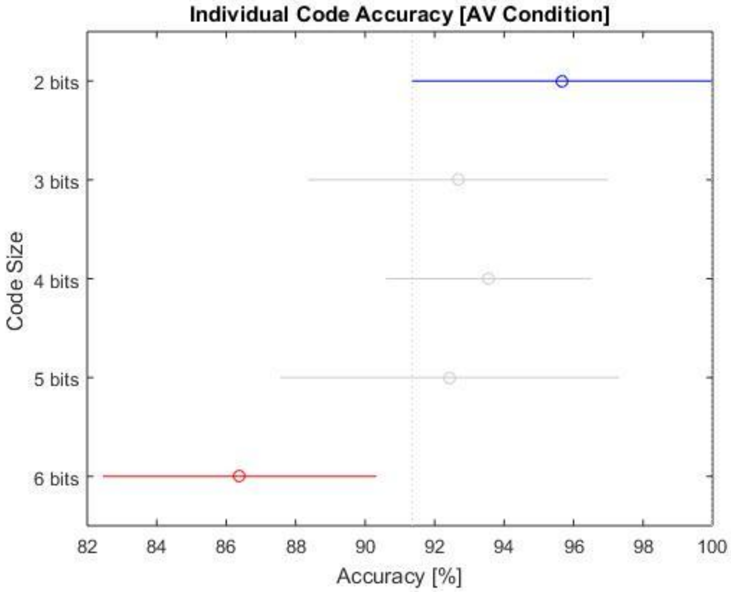


Figure 4.8: A pairwise comparison of the AV condition accuracy (%) for individual codes

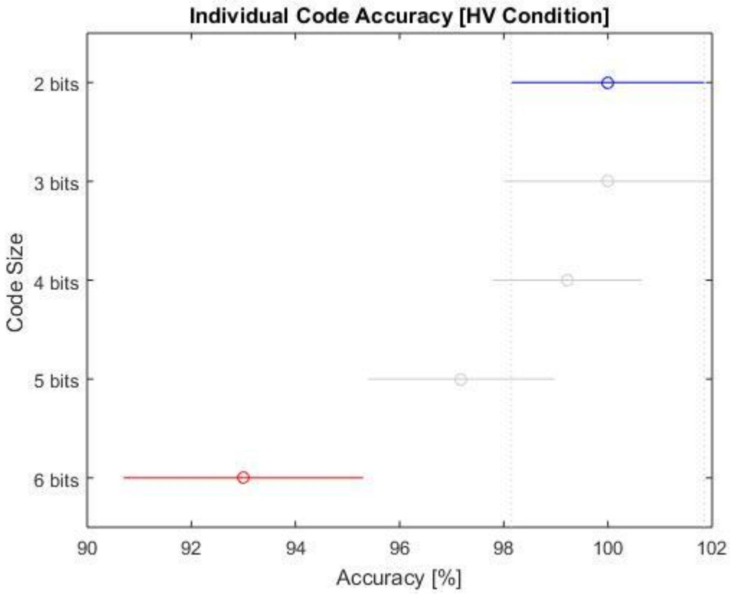


Figure 4.9: A pairwise comparison of the HV condition accuracy (%) for individual codes

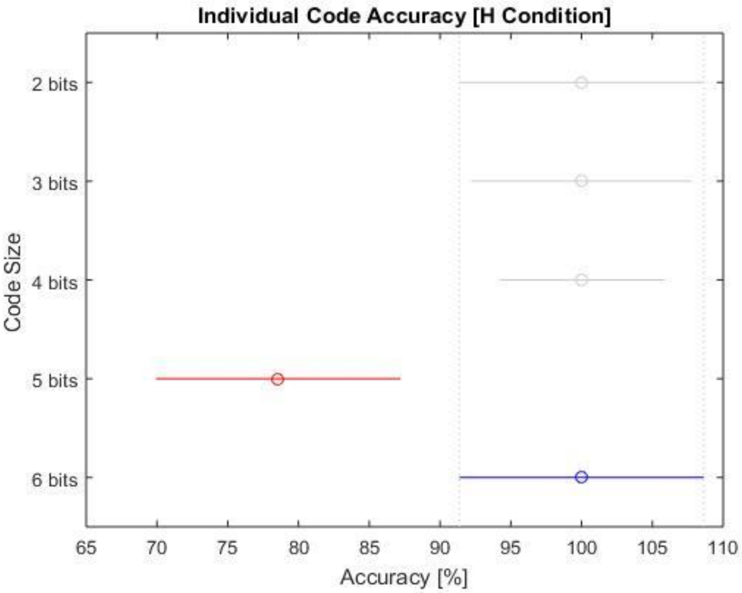


Figure 4.10: A pairwise comparison of the H condition accuracy (%) for individual codes

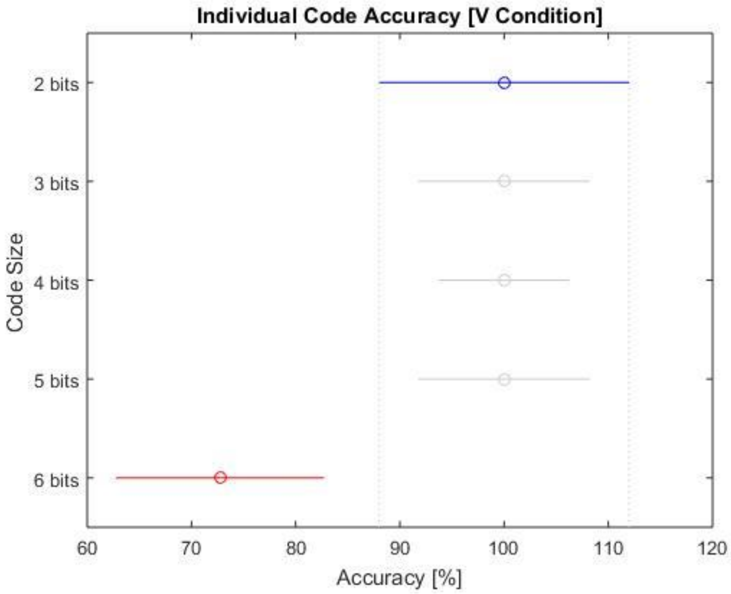


Figure 4.11: A pairwise comparison of the V condition accuracy (%) for individual codes

4.5 Conclusion

The impact of cognitive workload has shown some effect on participant's responses in the primary task. For instance, response time has slightly increased in Trial 2 for the AV condition and the HV condition. In addition, subjective performance has decreased in Trial 2 for all condition. Subjective workload is also affected by cognitive workload and can be seen by the significant increase in Trial 2 for all conditions. However, the impact is statistically not confirmed due to choosing codes that are less than the absolute judgement is one of the major problems in this study. Thus, to accurately determine whether there are differences between conditions, individual codes are analyzed. Additionally, 4 bits and 5 bits codes are not consistent with constant increase in response time or the constant decrease in accuracy as code size increment. This might be attributable to the fact that 4 bits and 5 bits may completely overlap with the primary task while 2 bits and 3 bits being shorter and 6 bits being longer do not completely overlap with the secondary task as shown in Figure 4.11. The overlap occurs when the combined feedback in the primary task coincide with the feedback from the secondary task.

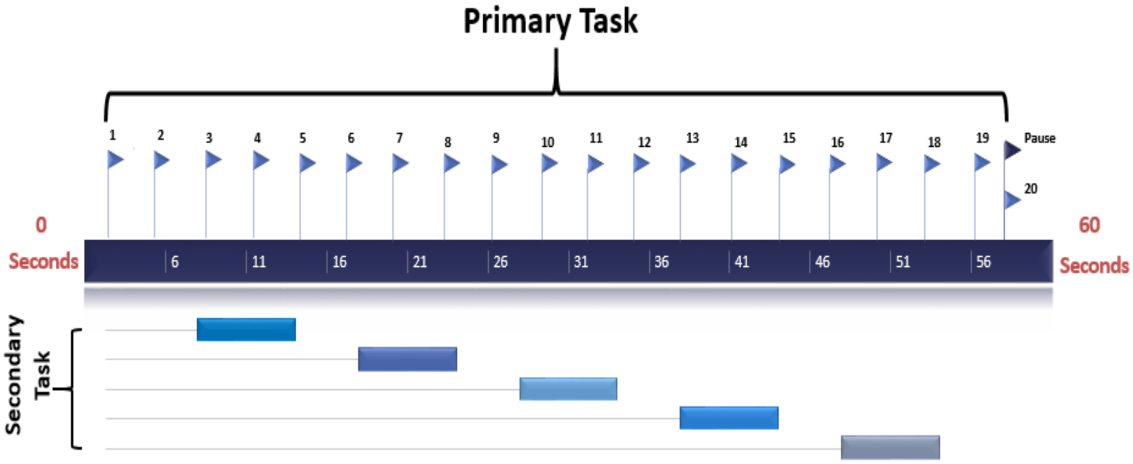


Figure 4.12: Primary and secondary task timeline for Trial 2 and Trial 3.

Chapter 5

Conclusion and Future Work

The purpose of this thesis is to examine the human performance and cognitive workload in combining AV, HV or AH senses in a primary task and H, A or V in a secondary task. The results of this study provide the same supporting principles for building system communication in which users collaborate effectively and efficiently in HAVE environments.

The results of the hypotheses tests are summarized below:

$H_{\Pr imary\ Task}$: The results show that the null hypothesis of the primary task can be rejected. The results indicate that participant's average response time for the audio-haptic (AH) condition is higher than the response time for the audio-visual (AV) and the haptic-visual (HV) conditions. In addition, the subjective performance for the HV condition is perceived to be higher than the AV and AH conditions with or without cognitive loading.

$H_{Secondary\ Task}$: There are several trends that can be seen from the results of adding a secondary task. Nevertheless, these trends cannot be confirmed from the ANOVA and multiple two-sample t-test hypothesis testing. The only confirmed difference is that the accuracy of the audio sensory modality is higher than the accuracy of the visual or the haptic sensory modality with the added cognitive loading of a secondary task.

The codes’ response time and accuracy are individually tested. The size of codes are between 2 bits and 6 bits. This is an inadequate size to pose a sufficient amount of workload on all participants. As a matter of fact, the code size is within what the human working memory can hold. According to Miller [51], the number of items that human working memory can hold is 5 items ± 2. The tendency for participants to hold the same amount of information can also be seen in the analysis of individual codes for the secondary task.

The lack of effective feedback is arguably one of the most common problems in the designing of interfaces. The results of this thesis provide the basic foundations on the design for human-in-the-loop applications. Most of the current implementations either utilize visual only or auditory only sensory modalities to provide feedback to users. Moreover, some studies has shown that only using one sensory modality is inadequate [23], [25]. To address these problems, sensory modality replacement or sensory modality addition to current and future interfaces may be adapted. This research study can potentially be used in HMI, tele-operation, collaboration, communication and medical applications.

In terms of tele-operation, the aviation industry is a promising area for multi-modal feedback to be implemented. For instance, the ground control station of an Unmanned Aerial Vehicles (UAVs) can also be enhanced with an added haptic feedback. While most commercial and military airplanes provide haptic feedback by nature (except some fly-by-wire airplanes) since they are mechanically operated, UAVs do not have haptic feedback. The addition of haptic feedback to a UAVs’ ground controller can enhance the user experience and increase the performance. It has also been proven that virtual forces increase the accuracy and decrease the time to complete the tele-operated task [3] in an assembly task.

The results of this research study supports the use of added haptic feedback without compromising primary task performance in the design of HMIs. Most of the warning in computers are visual warnings. As a result, the visual sensory modality of a computer user is already challenged. If the auditory sensory modality is being reserved for receiving auditory content, visual messages, such as "low battery", can distract the user. To solve this, messages can be haptically sent through available computer peripherals such as a mouse or a trackpad. Considering the limits of our senses, using one additional sensory feedback modality can be more suitable. Alternatively, visual feedback can be used where other types of feedback are indistinguishable. For example, car engine sound provides a lot of information about its condition. However, humans might not have the ability to distinguish between a healthy engine and a faulty engine. Therefore, a change in engine sound can be transmitted as visual information for the car operator to examine.

This thesis also studies the effects of cognitive loading on users' judgement and the ability of users to recognize an encoded message when they are engaged in another demanding task. The results of the study on workload can specifically be used on human bi-literal communication. In other words, the type of sensory modality used to communicate can, depending on the context, be chosen to effectively convey a message. For instance, a pilot crew is frequently engaged with multiple simultaneous communications; in a multi-crew setting, first officers communicate, which usually are in charge of communication and monitoring airplane parameters, with the pilot in command (PIC) and they also communicate with Air Traffic Control (ATC) using an auditory channel. Therefore, a haptic turn-taking protocol, similar to the turn-taking protocol defined in [65], can be used for pilots to relinquish control of the airplane and to acknowledge control of the airplane.

Another application of this study is the confirmation and testing of mild Traumatic Brain Injury (mTBI). With over 57 million mTBI incidents happening worldwide [67], it is very crucial to diagnose this medical phenomenon. Mild Traumatic Brain Injury (mTBI) also known as concussion is an injury caused by an instantaneous and aggressive shaking of the head [68]. Many methods have been developed to determine whether a person has a concussion or not and to evaluate the severity of this concussion. The current evaluation tools include questionnaires, physical diagnoses, self-reported checklists, or medical imaging such as MRI and CT scan [69]. These evaluation tools, however, have many disadvantages. First, questionnaires have a limited dependability especially when the injury is overestimated or under estimated, thus, a second evaluation tool is coupled with questionnaire. If two evaluation tools are used, then the evaluation process will take a long time to complete. Consequently, more reliable tests are used to evaluate concussion like MRI and CT scan. Despite the availability of highly reliable methods like MRI and CT scan, these methods are, unfortunately, very expensive and time-consuming. Additionally, electromagnetic radiation and x-ray radiation are potential risks in these test. This research study evaluates visual, auditory and haptic capabilities and motor response of a person. Thus, it can be a potential evaluation tool to determine whether a person is concussed or not. A test pre-concussion can be conducted as a baseline measure and a test post-concussion can be conducted to assess the cognitive abilities of patients after concussion.

In conclusion, significant research has to be done in the study of human performance and cognitive workload. In this thesis, human performance and cognitive workload is evaluated in a 2-dimensional level 2 collaboration [14]. Level 2 collaboration refers to a collaborative HAVE where one user has the ability to manipulate the virtual environment at a time. In the same context, level three collaboration should be studied by including another participant to the application over a network to ensure they are consistent. This will give a clear idea about the added benefits of level 3 collaboration over level 2 collaboration with respect to task performance and the distribution of resources to manage the cognitive workload. Additionally, the encoded messages should be extended beyond the human sensory modality absolute judgement to further study the effects of the secondary task on the response time and accuracy of the primary task. Moreover, the current application should be scaled to 3-dimensional graphical, auditory and haptic interfaces.

Appendices

Appendix A

Results

Participant	Condition	Response Time [sec]	Standard Deviation [sec]
1	AV	0.5859	0.3172
2	HV	0.6701	0.1389
3	AH	0.7021	0.3431
4	AV	0.8545	0.2281
5	HV	0.8683	0.0818
6	AH	0.9542	0.1046
7	AV	0.5147	0.1416
8	HV	0.6413	0.0755
9	AH	0.6487	0.1440
10	AV	0.7414	0.2349
11	HV	0.9528	0.1899
12	AH	0.8970	0.5538
13	AV	0.3616	0.2809
14	HV	0.7437	0.4621
15	AH	0.5946	0.2082
16	AV	0.6259	0.1229
17	HV	0.7292	0.1831
18	AH	0.7340	0.2823
19	AV	0.5939	0.1679
20	HV	0.6318	0.1083
21	AH	0.6031	0.3312
22	AV	0.7128	0.1861
23	HV	0.8397	0.2740
24	AH	0.7704	0.2817
25	AV	0.6438	0.2330
26	HV	0.7483	0.0451
27	AH	0.6324	0.1404
28	AV	0.6739	0.2446
29	HV	1.3234	0.2141

Table A.1: Response Time mean and standard deviation of primary task for each participant.

Participant	Condition	Accuracy [%]
1	AV	95.96
2	HV	82.76
3	AH	89.90
4	AV	98.99
5	HV	93.94
6	AH	94.95
7	AV	100.00
8	HV	100.00
9	AH	91.92
10	AV	97.98
11	HV	97.98
12	AH	92.93
13	AV	71.72
14	HV	97.98
15	AH	97.98
16	AV	97.98
17	HV	97.98
18	AH	98.99
19	AV	97.98
20	HV	100.00
21	AH	98.99
22	AV	95.96
23	HV	94.95
24	AH	90.91
25	AV	94.95
26	HV	96.97
27	AH	91.92
28	AV	98.99
29	HV	92.93

Table A.2: Accuracy of primary task for each participant.

Participant	Condition	Trial	Response Time [sec]	Standard Deviation [sec]
1	AV	2	36.2633	0.2766
		3	33.0842	0.6396
2	HV	2	13.6236	2.8614
		3	10.3332	0.1387
3	AH	2	35.7557	0.2628
		3	33.1172	8.7782
4	AV	2	52.5196	0.3885
		3	60.0000	0.5439
5	HV	2	36.0049	0.1293
		3	15.6257	8.1218
6	AH	2	60.0000	0.0883
		3	60.0000	0.0197
7	AV	2	50.2579	0.3434
		3	13.9879	0.3472
8	HV	2	36.1890	0.1892
		3	18.3975	0.1666
9	AH	2	24.3791	0.1292
		3	24.2579	0.1704
10	AV	2	60.0000	0.4631
		3	43.4787	0.4870
11	HV	2	60.0000	0.1961
		3	32.4754	1.0329
12	AH	2	60.0000	0.5625
		3	60.0000	0.3970
13	AV	2	23.6485	0.5537
		3	20.9104	0.2659
14	HV	2	46.1024	1.6139
		3	48.1728	0.2839

Participant	Condition	Trial	Response Time [sec]	Standard Deviation [sec]
15	AH	2	19.7655	0.1657
		3	15.3956	0.1950
16	AV	2	25.4044	0.2049
		3	25.5966	0.2146
17	HV	2	16.7854	0.1402
		3	17.0157	0.0978
18	AH	2	30.2724	0.4482
		3	26.5450	0.3808
19	AV	2	12.2047	0.1611
		3	15.7826	0.1278
20	HV	2	11.3811	0.1403
		3	11.4728	0.1361
21	AH	2	44.3913	0.4478
		3	20.3494	0.2420
22	AV	2	35.1066	0.4553
		3	38.4589	0.5919
23	HV	2	45.3088	0.1845
		3	55.7109	0.3266
24	AH	2	36.2767	0.2562
		3	31.8212	0.2267
25	AV	2	19.8273	0.3916
		3	16.5703	0.1938
26	HV	2	22.8401	0.1109
		3	16.7581	0.0728
27	AH	2	45.2911	0.3855
		3	16.5316	0.1387
28	AV	2	29.5728	1.2677
		3	30.3094	0.3102
29	HV	2	27.0655	0.3460
		3	18.3917	0.2591

Table A.3: Response time mean of secondary task for each participant.

Participant	Condition	Trial	Accuracy [%]
1	AV	2	60
		3	100
2	HV	2	100
		3	100
3	AH	2	80
		3	100
4	AV	2	80
		3	80
5	HV	2	100
		3	100
6	AH	2	100
		3	100
7	AV	2	100
		3	100
8	HV	2	100
		3	100
9	AH	2	100
		3	100
10	AV	2	80
		3	80
11	HV	2	100
		3	100
12	AH	2	40
		3	60
13	AV	2	100
		3	100
14	HV	2	80
		3	100
15	AH	2	100
		3	100

Participant	Condition	Trial	Accuracy [%]
16	AV	2	100
		3	100
17	HV	2	100
		3	100
18	AH	2	100
		3	100
19	AV	2	100
		3	100
20	HV	2	100
		3	100
21	AH	2	40
		3	100
22	AV	2	80
		3	60
23	HV	2	100
		3	100
24	AH	2	80
		3	100
25	AV	2	100
		3	100
26	HV	2	100
		3	100
27	AH	2	60
		3	100
28	AV	2	80
		3	100
29	HV	2	100
		3	100

Table A.4: Accuracy of secondary task for each participant.

Code Size [bits]	Total Number of Codes	Response Time [sec]	Standard Deviation [sec]
2	16	0.6447	0.3720
3	17	0.7297	0.3583
4	37	0.7869	0.6705
5	14	0.7749	0.4979
6	16	0.4715	4.5016

Table A.5: Individual code response time for primary task (AV condition).

Code Size [bits]	Total Number of Codes	Response Time [sec]	Standard Deviation [sec]
2	17	0.6275	0.1824
3	15	0.9064	4.3751
4	31	1.0170	8.8401
5	16	1.0095	1.6760
6	11	0.7024	0.2838

Table A.6: Individual code response time for primary task (HV condition).

Code Size [bits]	Total Number of Codes	Response Time [sec]	Standard Deviation [sec]
2	8	0.85893	0.26436
3	16	0.81913	0.28606
4	32	1.11016	4.79564
5	16	0.85082	0.32049
6	8	0.90214	0.47894

Table A.7: Individual code response time for primary task (AH condition).

Code Size [bits]	Total Number of Codes	Accuracy [%]
2	16	96.55
3	17	98.78
4	37	97.38
5	14	95.85
6	16	93.89

Table A.8: Individual code accuracy for primary task (AV condition).

Code Size [bits]	Total Number of Codes	Accuracy [%]
2	17	98.55
3	15	99.06
4	31	99.14
5	16	98.40
6	11	96.00

Table A.9: Individual code accuracy for primary task (HV condition).

Code Size [bits]	Total Number of Codes	Accuracy [%]
2	8	99.38
3	16	98.53
4	32	99.29
5	16	98.89
6	8	95.93

Table A.10: Individual code accuracy for primary task (AH condition).

Code Size [bits]	Total Number of Codes	Response Time [sec]	Standard Deviation [sec]
2	16	33.209	20.2443
3	17	31.6493	21.2123
4	37	46.2771	15.084
5	14	35.7699	18.8172
6	16	41.9882	18.2641

Table A.11: Individual code response time for secondary task (H condition).

Code Size [bits]	Total Number of Codes	Response Time [sec]	Standard Deviation [sec]
2	17	23.3798	21.1659
3	15	20.8922	21.1149
4	31	30.842	17.5282
5	16	22.4068	15.0751
6	11	34.5547	14.5297

Table A.12: Individual code response time for secondary task (A condition).

Code Size [bits]	Total Number of Codes	Response Time [sec]	Standard Deviation [sec]
2	8	26.9313	20.6977
3	16	26.6051	20.3038
4	32	39.0426	17.1986
5	16	32.7515	14.6141
6	8	51.2429	13.3289

Table A.13: Individual code response time for secondary task (V condition).

Code Size [bits]	Total Number of Codes	Accuracy [%]
2	16	87.50
3	17	100.00
4	37	91.89
5	14	78.57
6	16	87.50

Table A.14: Individual code accuracy for secondary task (H condition).

Code Size [bits]	Total Number of Codes	Accuracy [%]
2	17	100
3	15	100
4	31	100
5	16	100
6	11	100

Table A.15: Individual code accuracy for secondary task (A condition).

Code Size [bits]	Total Number of Codes	Accuracy [%]
2	8	100.00
3	16	94.12
4	32	88.89
5	16	88.89
6	8	72.73

Table A.16: Individual code accuracy for secondary task (V condition).

Participant	Condition	Trial	Perceived Performance	Perceived Workload
1	AV	1	12	40
		2	4	78
		3	11	52
2	HV	1	19	16
		2	17	33
		3	17	27
3	AH	1	19	22
		2	17	48
		3	17	44
4	AV	1	19	16
		2	17	33
		3	17	27
5	HV	1	19	22
		2	17	48
		3	17	44
6	AH	1	19	36
		2	13	59
		3	14	66
7	AV	1	19	22
		2	17	48
		3	17	44
8	HV	1	19	36
		2	13	59
		3	14	66
9	AH	1	6	35
		2	9	40
		3	6	43
10	AV	1	19	36
		2	13	59
		3	14	66

Participant	Condition	Trial	Perceived Performance	Perceived Workload
11	HV	1	6	35
		2	9	40
		3	6	43
12	AH	1	19	31
		2	18	56
		3	18	55
13	AV	1	6	35
		2	9	40
		3	6	43
14	HV	1	19	31
		2	18	56
		3	18	55
15	AH	1	19	23
		2	19	45
		3	19	33
16	AV	1	19	31
		2	18	56
		3	18	55
17	HV	1	19	23
		2	19	45
		3	19	33
18	AH	1	19	14
		2	7	43
		3	9	41
19	AV	1	19	23
		2	19	45
		3	19	33
20	HV	1	19	14
		2	7	43
		3	9	41
21	AH	1	10	48
		2	12	50
		3	7	58

Participant	Condition	Trial	Perceived Performance	Perceived Workload
22	AV	1	19	14
		2	7	43
		3	9	41
23	HV	1	10	48
		2	12	50
		3	7	58
24	AH	1	14	54
		2	11	76
		3	17	48
25	AV	1	10	48
		2	12	50
		3	7	58
26	HV	1	18	18
		2	16	37
		3	17	26
27	AH	1	19	24
		2	10	51
		3	19	17
28	AV	1	14	54
		2	11	76
		3	17	48
29	HV	1	16	37
		2	15	51
		3	16	45

Table A.17: Perceived performance and perceived workload for each participant.

Appendix B

ANOVA, Lilliefors and Two-Sample t-test Tables

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	0.39068	0.1953	5.58	0.0102
Error	24	0.8401	0.0350		
Total	26	1.2308			

Table B.1: Primary task response time ANOVA table for trial 1.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	0.2165	0.1082	5.09	0.0148
Error	23	0.4892	0.0212		
Total	25	0.7058			

Table B.2: Primary task response time ANOVA table for trial 2.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	0.3921	0.1960	2.17	0.1364
Error	24	2.1714	0.0904		
Total	26	2.5636			

Table B.3: Primary task response time ANOVA table for trial 3.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	22.771	11.3857	0.8	0.4616
Error	25	356.891	14.2756		
Total	27	379.662			

Table B.4: Primary task accuracy ANOVA table for trial 1.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	8.203	4.1014	0.89	0.4251
Error	23	106.228	4.6186		
Total	25	114.431			

Table B.5: Primary task accuracy ANOVA table for trial 2.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	66.74	33.3698	3.99	0.0309
Error	26	217.711	8.3735		
Total	28	284.45			

Table B.6: Primary task accuracy ANOVA table for trial 3.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	311.03	155.515	0.66	0.523
Error	26	6083.39	233.976		
Total	28	6394.42			

Table B.7: Secondary task response time ANOVA table for trial 2.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	285.97	142.986	0.66	0.527
Error	25	5438.66	217.547		
Total	27	5724.63			

Table B.8: Secondary task response time ANOVA table for trial 3.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	2251.85	1125.93	4.47	0.0224
Error	24	6044.44	251.85		
Total	26	8296.3			

Table B.9: Secondary task accuracy ANOVA table for trial 2.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	118.519	59.2593	2.29	0.1234
Error	24	622.222	25.9259		
Total	26	740.741			

Table B.10: Secondary task accuracy ANOVA table for trial 3.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	42.056	21.0282	1.46	0.2519
Error	25	360.622	14.4249		
Total	27	402.679			

Table B.11: Subjective performance ANOVA table for trial 1.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	138.101	69.0504	3.82	0.0353
Error	26	470.589	18.0996		
Total	28	608.69			

Table B.12: Subjective performance ANOVA table for trial 2.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	131.278	65.6389	6.23	0.0066
Error	24	252.722	10.5301		
Total	26	384			

Table B.13: Subjective performance ANOVA table for trial 3.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	237.99	118.997	0.5	0.6138
Error	26	6219.8	239.223		
Total	28	6457.79			

Table B.14: Subjective workload ANOVA table for trial 1.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	1195	597.498	2.19	0.1326
Error	26	7105.56	273.291		
Total	28	8300.55			

Table B.15: Subjective workload ANOVA table for trial 2.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Conditions	2	1206.41	513.204	2.95	0.0709
Error	25	4353.7	174.148		
Total	27	5380.11			

Table B.16: Subjective workload ANOVA table for trial 3.

Treatment		Test Statistics	Critical Value	P value
Response Time	AV [Trial 1]	0.1655	0.2740	0.6949
	AV [Trial 2]	0.1506	0.2740	0.8218
	AV [Trial 3]	0.2543	0.2740	0.0935
	HV [Trial 1]	0.1670	0.2740	0.6571
	HV [Trial 2]	0.3728	0.2740	7.08×10^{-4}
	HV [Trial 3]	0.2933	0.2740	0.02959
	AH [Trial 1]	0.1177	0.2740	0.9690
	AH [Trial 2]	0.1928	0.2740	0.4211
	AH [Trial 3]	0.2114	0.274	0.2910
Accuracy	AV [Trial 1]	0.4043	0.2740	1×10^{-3}
	AV [Trial 2]	0.2510	0.2740	0.1029
	AV [Trial 3]	0.2025	0.2740	0.3478
	HV [Trial 1]	0.2304	0.2740	0.1831
	HV [Trial 2]	0.2692	0.2740	0.0585
	HV [Trial 3]	0.3046	0.2740	0.0165
	AH [Trial 1]	0.2851	0.2740	0.0345
	AH [Trial 2]	0.3110	0.2740	0.0128
	AH [Trial 3]	0.2848	0.2740	0.0349

Table B.17: Primary task Lilliefors test of normality.

Treatment		Test Statistics	Critical Value	P value
Response Time	H [Trial 2]	0.1664	0.2740	0.6687
	H [Trial 3]	0.1620	0.2740	0.7046
	A [Trial 2]	0.1516	0.2740	0.7964
	A [Trial 3]	0.3239	0.2740	0.0075
	V [Trial 2]	0.1469	0.2740	0.8288
	V [Trial 3]	0.2516	0.274	0.1011
Accuracy	H [Trial 2]	0.3333	0.2740	0.0050
	H [Trial 3]	0.3963	0.2740	8.33×10^{-4}
	A [Trial 2]	0.5194	0.2740	1.00×10^{-4}
	A [Trial 3]	1	0.2740	1.00×10^{-5}
	V [Trial 2]	0.2538	0.2740	0.1060
	V [Trial 3]	0.5194	0.2740	1.00×10^{-4}

Table B.18: Secondary task Lilliefors test of normality.

Treatment		Test Statistics	Critical Value	p value
Response Time	AV [Trial 1]	0.4024	0.2740	1.00×10^{-4}
	AV + H [Trial 2]	0.2244	0.2740	0.2124
	AV + H [Trial 3]	0.2267	0.2740	0.1950
	HV [Tail 1]	0.2430	0.2740	0.1368
	HV + A [Trial 2]	0.2177	0.2740	0.2378
	HV + A [Trial 3]	0.2079	0.2740	0.3193
	AH [Trial 1]	0.2409	0.2740	0.1431
	AH + V [Trial 2]	0.2222	0.2740	0.2307
	AH + V [Trial 3]	0.2738	0.2740	0.0370
Accuracy	AV [Trial 1]	0.1577	0.2740	0.7606
	AV + H [Trial 2]	0.1734	0.2740	0.6096
	AV + H [Trial 3]	0.1373	0.2740	0.8930
	HV [Tail 1]	0.1898	0.2740	0.4458
	HV + A [Trial 2]	0.1514	0.2740	0.7936
	HV + A [Trial 3]	0.1649	0.2740	0.6901
	AH [Trial 1]	0.1828	0.2740	0.5273
	AH + V [Trial 2]	0.1909	0.2740	0.4573
	AH + V [Trial 3]	0.1942	0.2740	0.4226

Table B.19: Subjective performance and subjective workload Lilliefors test of normality.

Conditions			df	Confidence Interval		Standard Deviation	t-statistics	p-value
				Lower Bound	Higher Bound			
Response Time	AV [Trial 1]	HV [Trial 1]	16	-0.1308	0.1308	0.1539	0.3287	0.7466
	AV [Trial 1]	AH [Trial 1]	15	-0.4298	-0.0522	0.1823	-2.7204	0.0158
	HV [Trial 1]	AH [Trial 1]	17	-0.4756	-0.0544	0.2173	-2.6545	0.0167
	AV [Trial 2]	HV [Trial 2]	15	-0.0417	0.1930	0.1133	1.3744	0.1895
	AV [Trial 2]	AH [Trial 2]	16	-0.3096	0.0193	0.1646	-1.8714	0.0797
	HV [Trial 2]	AH [Trial 2]	15	-0.3797	-0.0620	0.1534	-2.9634	0.0097
	AV [Trial 3]	HV [Trial 3]	16	-0.4028	0.1511	0.2771	-0.9633	0.3497
	AV [Trial 3]	AH [Trial 3]	16	-0.5541	-0.0343	0.2601	-2.3994	0.0290
	HV [Trial 3]	AH [Trial 3]	16	-0.5244	0.1878	0.3564	-1.0002	0.3313
Accuracy	AV [Trial 1]	HV [Trial 1]	17	-1.9953	1.5240	1.8152	-0.2826	0.7809
	AV [Trial 1]	AH [Trial 1]	16	-2.6966	6.2880	4.4953	0.8474	0.4093
	HV [Trial 1]	AH [Trial 1]	16	-2.2683	6.3311	4.4354	0.8474	0.3328
	AV [Trial 2]	HV [Trial 2]	17	-2.9619	1.7497	2.4302	-0.5428	0.5943
	AV [Trial 2]	AH [Trial 2]	14	-3.6210	0.7350	2.0150	-1.4210	0.1772
	HV [Trial 2]	AH [Trial 2]	15	-2.8522	1.1783	1.9186	-0.8852	0.3900
	AV [Trial 3]	HV [Trial 3]	18	-6.2546	-0.5987	3.0099	-2.5457	0.0203
	AV [Trial 3]	AH [Trial 3]	17	-5.7974	0.0960	3.0398	-2.0411	0.0571
	HV [Trial 3]	AH [Trial 3]	17	-1.9484	3.1002	2.6040	0.4814	0.6364

Table B.20: Primary task multiple two-sample t-test.

Conditions			df	Confidence Interval		Standard Deviation	t-statistics	p-value
				Lower Bound	Higher Bound			
Response Time	H [Trial 2]	A [Trial 2]	18	-11.8356	17.7364	15.7371	0.4192	0.6800
	H [Trial 2]	V [Trial 2]	17	-19.5629	9.3834	14.9301	-0.7420	0.4682
	A [Trial 2]	V [Trial 2]	17	-22.7596	6.6793	15.1842	-1.1524	0.2651
	H [Trial 3]	A [Trial 3]	17	-11.3413	15.1805	13.5668	0.3255	0.7488
	H [Trial 3]	V [Trial 3]	16	-19.6724	8.5970	14.1441	-0.8305	0.4185
	A [Trial 3]	V [Trial 3]	17	-23.4213	8.2881	16.3553	-1.0069	0.3281
Accuracy	H [Trial 2]	A [Trial 2]	16	-16.3375	-1.4403	7.4536	-2.5298	0.0223
	H [Trial 2]	V [Trial 2]	16	-6.0902	32.7569	19.4365	1.4552	0.1649
	A [Trial 2]	V [Trial 2]	16	4.2836	40.1608	17.9505	2.6261	0.0183
	H [Trial 3]	A [Trial 3]	17	-10.3092	1.4203	6.0499	-1.5989	0.1283
	H [Trial 3]	V [Trial 3]	15	-11.1150	2.2261	6.4406	-1.4201	0.1760
	A [Trial 3]	V [Trial 3]	16	-4.9774	13.8662	9.4281	1.0000	0.3322

Table B.21: Secondary task multiple two-sample t-test.

Conditions			df	Confidence Interval		Standard Deviation	t-statistics	p-value
				Lower Bound	Higher Bound			
Subjective Performance	AV [Trial 1]	HV [Trial 1]	17	-4.4792	1.0126	2.8326	-1.3318	0.2005
	AV [Trial 1]	AH [Trial 1]	16	-3.3178	5.7622	4.5430	0.5707	0.5761
	HV [Trial 1]	AH [Trial 1]	17	-0.7961	6.7072	3.8701	1.6621	0.1148
	AV + H [Trial 2]	HV + A [Trial 2]	18	-8.6962	-1.5038	3.8275	-2.9795	0.0080
	AV + H [Trial 2]	AH + V [Trial 2]	17	-6.3260	3.5037	5.0700	-0.6058	0.5527
	HV + A [Trial 2]	AH + V [Trial 2]	17	0.0428	7.3349	3.7612	2.1346	0.0476
	AV + H [Trial 3]	HV + A [Trial 3]	17	-8.5157	-1.3732	3.6840	-2.9211	0.0095
	AV + H [Trial 3]	AH + V [Trial 3]	16	-7.8496	-0.1504	3.8283	-2.2027	0.0426
	HV + A [Trial 3]	AH + V [Trial 3]	15	-7.8496	-0.1504	3.8283	-2.2027	0.0426
Subjective Workload	AV [Trial 1]	HV [Trial 1]	18	-6.7514	16.7514	12.5073	0.8939	0.3832
	AV [Trial 1]	AH [Trial 1]	18	-18.3831	14.8498	17.1411	-0.2243	0.8252
	HV [Trial 1]	AH [Trial 1]	17	-22.7628	9.2294	16.5012	-0.8925	0.3846
	AV + H [Trial 2]	HV + A [Trial 2]	18	-1.5731	25.9731	14.6591	1.8610	0.0792
	AV + H [Trial 2]	AH + V [Trial 2]	17	-19.3816	14.5371	17.4948	-0.3013	0.7668
	HV + A [Trial 2]	AH + V [Trial 2]	17	-31.4783	2.2338	17.3882	-1.8302	0.0848
	AV + H [Trial 3]	HV + A [Trial 3]	18	1.8480	25.9520	12.8273	2.4231	0.0262
	AV + H [Trial 3]	AH + V [Trial 3]	16	-2.6679	23.0679	12.7967	1.6804	0.1123
	HV + A [Trial 3]	AH + V [Trial 3]	16	-17.7545	10.3545	13.9768	-0.5581	0.5845

Table B.22: Subjective performance and subjective workload two-sample t-test.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Sample	2	0.354355	0.177178	2.396178	0.101879
Columns	1	0.274481	0.274481	3.712122	0.05995
Interaction	2	0.193181	0.096591	1.306307	0.280272
Within	48	3.549206	0.073942		
Total	53				

Table B.23: Two-way ANOVA table for the response time of the primary task between Trial 1 and Trial 2.

Source	DF	Sum of Squares	Mean Square	f observed	P value
Sample	2	0.6787	0.33936	2.87	0.0627
Columns	1	0.5922	0.59108	4.99	0.0282
Interaction	2	0.3021	0.15106	1.28	0.2848
Within	81	9.5908	0.1184		
Total	86				

Table B.24: Two-way ANOVA table for the response time of the primary task between Trial 1 and the pooled data from Trial 2 and Trial 3.

Appendix C

Ethics Clearance and Questionnaires

By signing this consent form, you are not waiving your legal rights or releasing the investigator(s) or involved institution(s) from their legal and professional responsibilities.

Consent Form

I have read the information presented in the information letter about a study being conducted by Mortaja Alqassab under the supervision of Professor David Wang of the Department of Electrical and Computer Engineering at the University of Waterloo. I have had the opportunity to ask any questions about the study. I am aware that I may withdraw from the study without penalty at any time by advising the researcher of my decision.

I understand that this project has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Committee. I was informed that if I have any questions, comments or concerns resulting from my participation in this study, I may contact Maureen Nummelin, Office of Research Ethics Director, at 519-888-4567, Ext. 36005 or maureen.nummelin@uwaterloo.ca.

With full knowledge of all foregoing, I agree, of my own free will, to participate in this study.

Participant Name: ________________________________

Participant Signature: _____________________________

Witness Name: __________________________________

Witness Signature: ______________________________

Dated: _________________

March 16, 2015

Title of Study:	Human Performance and Cognitive Workload in Haptic Audio Visual Environments
Student Investigator:	Mortaja Alqassab, Electrical and Computer Engineering Department, malqassa@uwaterloo.ca
Department Supervisor:	David Wang, Electrical and Computer Engineering, dwang@uwaterloo.ca

Purpose the Study

I am conducting a research study as part of my M.A.Sc Thesis. The purpose of this research study is to investigate human performance and mental workload in tasks involving different types of feedback modes, namely haptic feedback, audio feedback, visual feedback or a combination of any feedback mode and to investigate the effects of different levels of workload on task completion and user performance. This study aims to determine the optimum feedback combination for complex human-machine or human-computer interaction.

Study Procedures

Participants in this study will be asked to interact in a virtual reality setting using a computer screen, a noise cancelling headset (Figure 2). A haptic device (Figure 1) is an instrument that sends small forces to the user through a pen like mechanism which simulates touching objects in the virtual world. The Bose A20 aviation headset is a noise cancelling headset which will be used to isolate outside noises. This study will take place in E5 5001 or EIT 3111.

After providing consent, you will be given some time to ask questions. After all question are answered, you will be given some time to read the test procedure. You will be asked to complete 3 virtual reality settings. Each setting will last 5 minutes. A one minute training session will precede each virtual reality setting. A break will be given after completing 3 tasks. After completing all virtual reality setting, you will be asked to fill out a questionnaire and asked general questions about your experience and you will be given a feedback and appreciation letter. This study will approximately last 30 minutes. Below is an outline for what you will be required to do:

1. For each virtual reality setting, you will be required to press a virtual buttons on the screen using the haptic device which is presented by a cursor on the screen (cone shaped) as shown in figure 3 after provided by one or a combination of the following indications:
   a. A flashing light from on the buttons on the screen.
   b. A tone (left ear, right ear, both ears) that indicates the direction of the button.
   c. A force from the haptic device (force to the right, force to the left, force in the up direction) that indicates the direction of the button.
2. For each virtual reality setting, you will be required to write a code on a paper (Morse code). The code is conveyed to you using one of the following methods:
   a. You feel vibrations on the haptic device.
   b. You hear tones from the headset.
   c. You observe flashing light on the screen.
3. For each virtual reality setting, you will be asked to answer a questionnaire.
   a. The NASA TLX Questionnaire consists of two parts.
   b. The first part of the NASA TLX requires the participant to answer 6 questions about demand, performance, effort and frustration with 20-point scale.
   c. The second part of the questionnaire requires the participant to choose which measurement is relevant to workload by providing the definition for each question.

Figure C.1: PHANToM Omni haptic device from Geomagic (formerly SensAble Technologies).

Figure C.2: Bose A20 aviation noise cancelling headset.

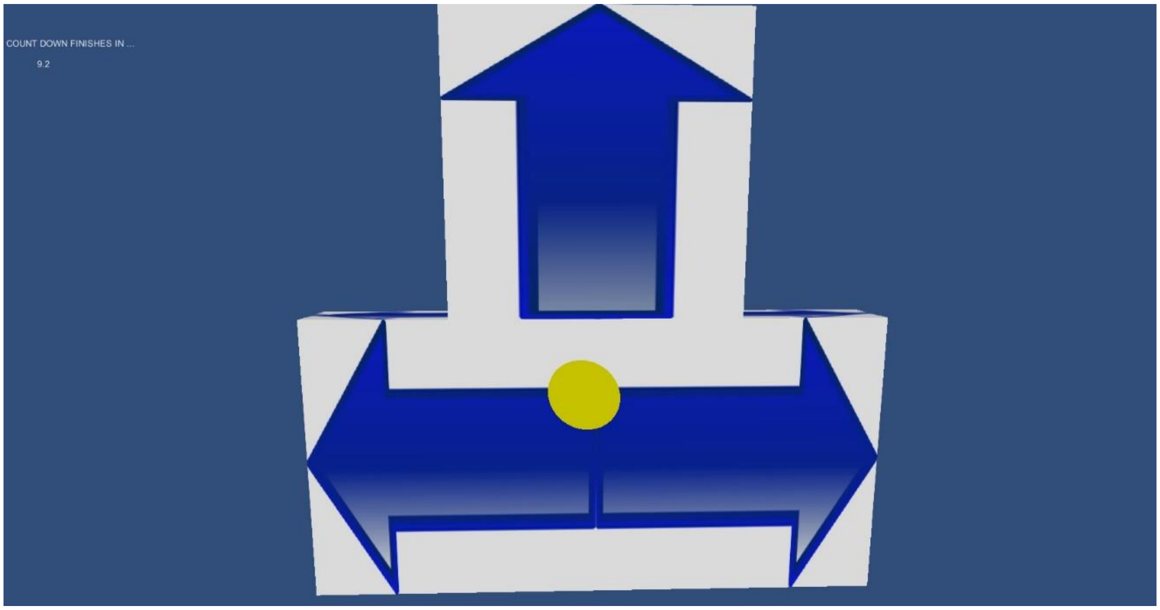


Figure C.3: Graphical User Interface (GUI) of the experiment.

Risks

There are not any anticipated risks in participating in this study. All devices are approved for commercial use and does not have any side effects on you as a participant. All equipment will be sanitized with alcohol wipes and student investigator will wash their hands and use a sanitizer after each session.

Potential Benefits

There are not any anticipated direct benefits to you as a participant. This study, however, will contribute to the research community by determining the most suitable feedback combination for collaborative and remote communication.

Withdrawal from the Study

Your participation in this study is completely voluntary. You may decide to withdraw from the study or refuse to answer any of the questions in whole or in part, at any time for any reason.

Confidentiality and Data Security

All information you provide will remain confidential; your name will not be included with the data collected in the study. You will not be identified individually in any written reports in this research as the interest is on the average responses of participants. The unidentified information or average responses of participants will be published in M.A.Sc thesis and/or in journal articles. All paper records of data collected in this study will be retained in a locked cabinet, in which only the researchers have access. All records will safely be disposed after 5 years from the completion of the study.

Questions and Concerns

If you have any question about the participation on the study, please do not hesitate to contact Prof. David Wang from the Department of Electrical and Computer Engineering at dwang@uwaterloo.ca mailto:c4burns@uwaterloo.caor 519-888-4567 x33968. I would like to assure you that the study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Committee. Should you have any questions about the study, contact Maureen Nummelin, Office of Research Ethics Director, at 519-888-4567, Ext. 36005 or maureen.nummelin@uwaterloo.ca.

UNIVERSITY OF
WATERLOO

Department of Electrical and Computer
Engineering

Standard Operating Procedures

Title of SOP:	*Procedures of the Use of Virtual Reality Interfaces in the Studies involving Collaborative Haptic, Audio and Visual Environments.*
Responsibility:	*Student Investigator.*
SOP created on:	*March 24, 2015*
SOP created by:	*Mortaja Alqassab*
Signature:	
Date:	

☐ I acknowledge that as the principal investigator/faculty supervisor I am responsible for updating this SOP and notifying the ORE through a modification form (Form 104) if any of the procedures as outlined above change or require revision.

A. Purpose of the Study Protocol

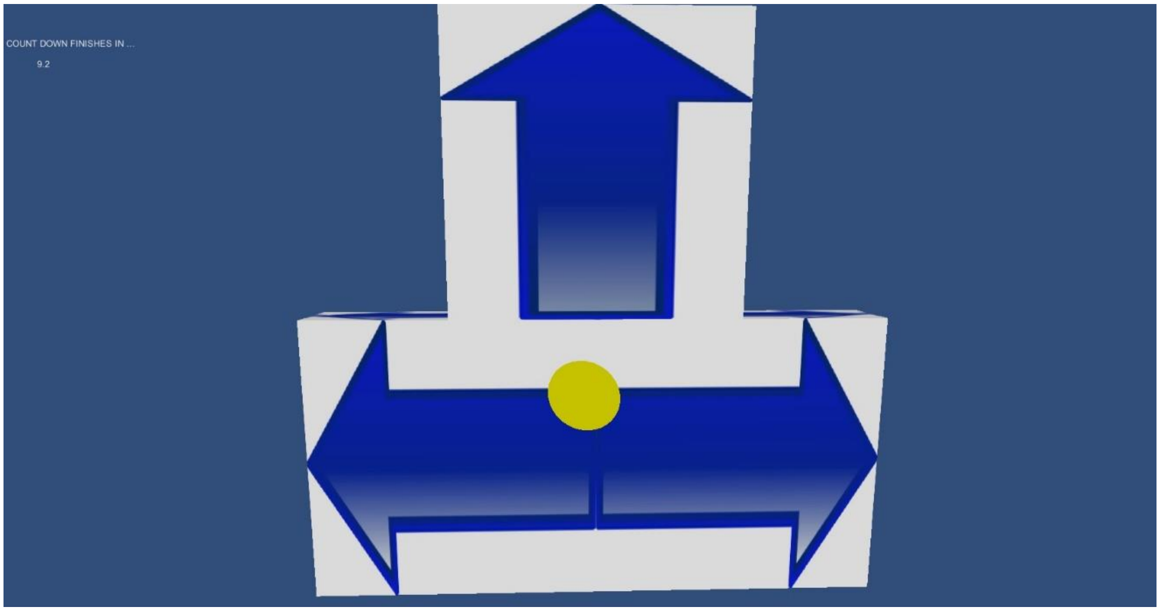
1. The study protocol outlines the steps the student investigator/research assistants should follow to achieve research tasks.
2. The study protocol outlines the equipment to be used in this study.
3. The study protocol outlines the safeguards the student investigator/research assistants should use to safely complete the research tasks.

B. Equipment

4. Participants are asked to use a haptic device to manipulate and interact with virtual setting. The haptic device is (PHANToM Omni) is manufactured by Geomagic (previously Sensable Technologies). The device is used to send and receive forces and position information produced by the user or by the computer virtual reality simulation.
5. Participants will be asked to wear noise cancelling headset to listen to audio.
6. Participants will be asked to use a computer monitor to observe a virtual reality setting.

Figure C.4: PHANToM Omni haptic device from Geomagic (formerly SensAble Technologies).

Figure C.5: Bose A20 aviaiton noise cancelling headset.



COUNT DOWN FINISHES IN ...

9.2

Figure C.6: Graphical User Interface (GUI) of the experiment.

C. Recruitment Procedure

1. Potential Participants are recruited into study using Office of Research Ethics accepted Procedures.
2. Participants will be asked if they are allergic to alcohol sanitizer. Participants with allergies to alcohol sanitizer will be excluded from study.
3. Potential participants will be asked to come to the lab on the scheduled date and time. The study will be explained to them by the student investigator. The participants will asked to read the information letter and will be given time to ask questions. After given satisfactory information about the study they will indicate their agreement by the signing the consent form. The participants will be given a copy of the information letter and consent form.

D. Study Procedures

1. The participants will be given some time to read the information letter and ask questions.
2. Participants will be given some time to ask questions.
3. After all questions are answered, participants will be asked to sign the consent form.
4. Participants will be informed that all equipment is sanitized by alcohol wipes as part of the study procedure. Participants will be asked if they have allergies to alcohol. Participants with alcohol allergies will be excused from the study.
5. Participants will be asked to wear the virtual reality equipment previously mentioned.
6. Participants will be asked to complete 3 virtual reality setting.
7. For each virtual reality setting, participants will be required to press virtual buttons on the screen using the haptic device which is presented by a cursor on the screen (cone shaped) after provided by the one or a combination of the following indications:
    a. A flashing light from on the buttons on the screen.
    b. A tone (left ear, right ear, both ears) that indicates the direction of the button.
    c. A force from the haptic device (force to the right, force to the left, force in the up direction) that indicates the direction of the button.
8. For each virtual reality setting, participants will be required to write a code on a paper (Morse code). The code is conveyed to participants using one of the following methods:
    a. Participants feel vibrations on the haptic device.
    b. Participants hear tones from the headset.
    c. Participants observe flashing light on the screen.
9. For each virtual reality setting participants will be asked to answer a

questionnaire.

   a. The NASA TLX Questionnaire consists of two parts.
   b. The first part of the NASA TLX requires the participant to answer 6 questions about demand, performance, effort and frustration with 20-point scale.
   c. The second part of the questionnaire requires the participant to choose which measurement is relevant to workload by providing the definition for each question.

10. Virtual Reality equipment is sanitized with alcohol wipes.

11. The student investigators wash their hands and use a sanitizer after each session.

E. Risks

1. **Participants**
   - **a.** There are not perceive risks to participants.
2. **Researchers**
   - **a.** There are not perceived risks to researchers.

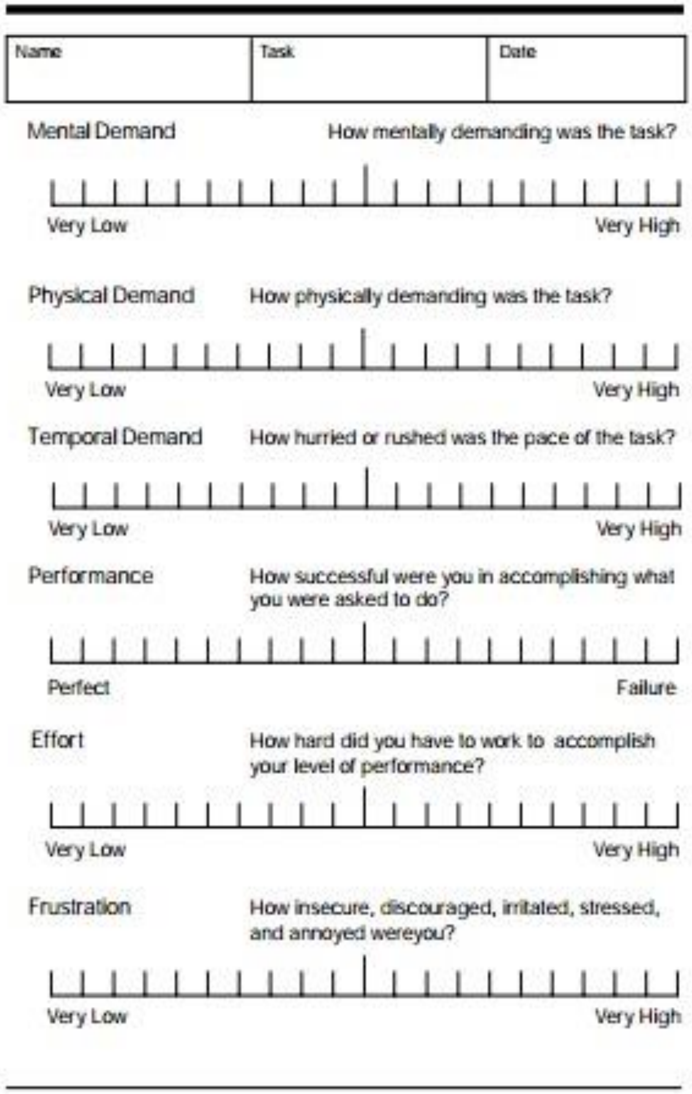
F. Safeguards/Safety Procedures

1. Virtual Reality equipment is sanitized with alcohol wipes.
2. The student investigators wash their hands and use a sanitizer after each session.

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.



Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Code Recognition

Participant: Task: Date:

Fill in the Codes in the spaces Avialable. Type “S” or a “•” for short vibrations, tones or flashes and type “L” or “–” for long vibrations, tones or flashes:

1.

2.

3.

4.

5.

References

[1] E. Sallnäs, K. Rassmus-Grohn, and C. Sjostrom, “Supporting Presence in Collaborative Environments by Haptic Force Feedback,” *ACM Trans. Comput. Interact.*, vol. 7, no. 4, pp. 461–476, 2000.

[2] J. N. Bailenson, N. Yee, S. Brave, D. Merget, and D. Koslow, “Virtual interpersonal touch: expressing and recognizing emotions through haptic devices,” *Human-Computer Interact.*, vol. 22, pp. 325–353, 2007.

[3] H. Boessenkool, D. a. Abbink, C. J. M. Heemskerk, and F. C. T. van der Helm, “Haptic shared control improves tele-operated task performance towards performance in direct control,” *2011 IEEE World Haptics Conf.*, pp. 433–438, Jun. 2011.

[4] S. Ullah, P. Richard, S. Otmane, M. Naud, and M. Mallem, “Haptic guides in cooperative virtual environments: Design and human performance evaluation,” *2010 IEEE Haptics Symp. HAPTICS 2010*, pp. 457–462, 2010.

[5] A. Chellali and I. Milleville, “Haptic communication to enhance collaboration in virtual Environments,” *Proc. Eur. Conf. Cogn. Ergon. ECCE2010*, pp. 83–90, 2010.

[6] X. S. X. Shen, J. Z. J. Zhou, a. El Saddik, and N. D. Georganas, “Architecture and Evaluation of Tele-Haptic Environments,” *Eighth IEEE Int. Symp. Distrib. Simul. Real-Time Appl.*, 2004.

[7] M. McLaughlin, G. Sukhatme, W. P. W. Peng, W. Z. W. Zhu, and J. Parks, “Performance and co-presence in heterogeneous haptic collaboration,” *11th Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst. 2003. HAPTICS 2003. Proceedings.*, 2003.

[8] M. Glencross, C. Jay, J. Feasel, L. Kohli, M. Whitton, and R. Hubbold, “Effective cooperative haptic interaction over the Internet,” *Proc. - IEEE Virtual Real.*, pp. 115–122, 2007.

[9] J. Kjölberg and E. Sallnäs, “Supporting Object Handling and Hand Over Tasks in Haptic Collaborative Virtual Environments,” *EuroHaptics 2002 Conf. Proc.*, pp. 71–76, 2002.

[10] G. A. Boy, *The Handbook of Human-Machine Interaction: A Human-Centered Design Approach*. Ashgate Publishing, Ltd., 2012.

[11] A. Magazine, “Angle of Attack,” 2011. [Online]. Available: http://www.boeing.com/commercial/aeromagazine/aero_12/attack_story.html. [Accessed: 30-Mar-2015].

[12] GMC, “2014 Sierra Safety Alert Seat Aids Driver Awareness.” [Online]. Available: http://media.gmc.com/media/us/en/gmc/news.detail.html/content/Pages/news/us/en/2013/Apr/0430-sierra.html. [Accessed: 09-Apr-2015].

[13] M. S. Pinho, D. a. Bowman, and C. M. D. S. Freitas, “Cooperative object manipulation in immersive virtual environments,” *Proc. ACM Symp. Virtual Real. Softw. Technol. - VRST ’02*, p. 171, 2002.

[14] D. Margery, B. Arnaldi, and N. Plouzeau, “A general framework for cooperative manipulation in virtual environments,” *Virtual Environ.*, vol. 99, pp. 169–178, 1999.

[15] B. Collins-Sussman, “The subversion project: buiding a better CVS,” *Linux J.*, vol. 2002, no. 94, p. 3, Feb. 2002.

[16] S. Ullah, X. Liu, S. Otmane, P. Richard, M. Mallem, S. Ullah, X. Liu, S. Otmane, P. Richard, M. Mallem, and W. You, “What You Feel Is What I Do : A Study of Dynamic Haptic Interaction in Distributed Collaborative Virtual Environment To cite this version :,” *Human-Computer Interact. Interact. Tech. Environ.*, no. 140–147, 2014.

[17] S. Ullah, P. Richard, S. Otmane, M. Naud, and M. Mallem, “Haptic Guides in Cooperative Virtual Environments: Design and Human Performance Evaluation,” *Haptics Symp. 2010 IEEE*, 2013.

[18] A. Girard, Y. Bellik, M. Auvray, and M. Ammi, “Visuo-Haptic tool for collaborative adjustment of selections,” *Haptic Audio Interation Des.*, 2013.

[19] A. Chellali, I. Milleville-pennel, A. Chellali, and I. M. Influence, “Influence of Haptic Communication on a Shared Manual Task in a Collaborative Virtual Environment.,” *Interact. Comput.*, vol. 23, no. 4, 2013.

[20] S. O. Oguz, A. Kucukyilmaz, T. M. Sezgin, and C. Basdogan, “Haptic negotiation and role exchange for collaboration in virtual environments,” *2010 IEEE Haptics Symp. HAPTICS 2010*, pp. 371–378, 2010.

[21] E. Sallnäs and S. Zhai, “Collaboration meets Fitts ’ law : Passing Virtual Objects with and without Haptic Force Feedback,” *INTERACT*, pp. 97–104, 2003.

[22] S. Ullah, X. Liu, S. Otmane, P. Richard, and M. Mallem, “What You Feel Is What I Do: A Study of Dynamic Haptic Interaction in Distributed Collaborative Virtual Environment,” *Human-Computer Interact. Interact. Tech. Environ. Pt Ii*, vol. 6762, pp. 140–147, 2011.

[23] Oakley, Brewster, and Gray, “Can you feel the force? An investigation of haptic collaboration in shared editors,” *Proc. EuroHaptics*, 2001.

[24] J. Moll and E. L. Sallnäs, “Communicative functions of haptic feedback,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5763 LNCS, pp. 1–10, 2009.

[25] J. Wang, A. Chellali, and C. G. L. Cao, “A study of communication modalities in a virtual collaborative task,” *Proc. - 2013 IEEE Int. Conf. Syst. Man, Cybern. SMC 2013*, pp. 542–546, 2013.

[26] S. Yantis, *Visual Perception: Essential Readings*. Psychology Press, 2001.

[27] H. Murase and S. Nayar, “Visual learning and recognition of 3-D objects from appearance,” *Int. J. Comput. Vis.*, 1995.

[28] M. Peterson, “Object recognition processes can and do operate before figure-ground organization,” *Curr. Dir. Psychol. Sci.*, 1994.

[29] M. Tarr, “Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects,” *Psychon. Bull. Rev.*, 1995.

[30] K. R. Boff and L. Kaufman, *Handbook of Perception and Human Performance: Sensory Processes and Perception*. Wiley, 1986.

[31] M. Born, E. Wolf, and A. B. Bhatia, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 1999.

[32] E. H. Weber, H. E. Ross, and D. J. Murray, *E.H. Weber on the Tactile Senses*. Psychology Press, 1996.

[33] R. L. Klatsky and S. J. Lederman, “Haptic Perception,” *Encyclopedia of Cognitive Science*. pp. 508–512, 2003.

[34] M. Grunwald, *Human haptic perception: Basics and applications*. 2008.

[35] M. H. Zadeh, D. Wang, and E. Kubica, “Human Factors for Designing a Haptic Interface for Interaction with a Virtual Environments,” *IEEE Int. Work. Haptic Audio Vis. Environ. their Appl.*, no. October, pp. 12–14, 2007.

[36] Y. Y. Huang, J. Moll, E. L. Sallnäs, and Y. Sundblad, “Auditory feedback in haptic collaborative interfaces,” *Int. J. Hum. Comput. Stud.*, vol. 70, no. 4, pp. 257–270, 2012.

[37] R. J. Teather, D. Natapov, and F. Law, “Evaluating Haptic Feedback in Virtual Environments using ISO 9241-9,” *Virtual Real. Conf. (VR), 2010 IEEE,* pp. 9–10, 2010.

[38] M. A. Srinivasan, G. L. Beauregard, and D. L. Brock, “The Impact of Visual Information on the Haptic Perception of Stiffness in Virtual Environments,” in *Proceedings of the ASME Dynamics Systems and Control Division*, 1996.

[39] K. Bark, P. Khanna, R. Irwin, P. Kapur, S. a. Jax, L. J. Buxbaum, and K. J. Kuchenbecker, “Lessons in using vibrotactile feedback to guide fast arm motions,” *2011 IEEE World Haptics Conf.*, pp. 355–360, Jun. 2011.

[40] J. J. Young, C. Stolfi, H. Z. Tan, J. Chevrier, B. Dick, and G. Bertoline, “Learning force concepts using visual trajectory and haptic force information at the elementary school level,” *2011 IEEE World Haptics Conf.*, pp. 391–396, Jun. 2011.

[41] M. Mulder, D. a. Abbink, and E. R. Boer, “The effect of haptic guidance on curve negotiation behavior of young, experienced drivers,” *2008 IEEE Int. Conf. Syst. Man Cybern.*, pp. 804–809, Oct. 2008.

[42] D. Hecht and M. Reiner, “Sensory dominance in combinations of audio, visual and haptic stimuli,” *Exp. Brain Res.*, vol. 193, no. 2, pp. 307–314, 2009.

[43] M. a. Zahariev and C. L. MacKenzie, “Auditory, graphical and haptic contact cues for a reach, grasp, and place task in an augmented environment,” *Proc. 5th Int. Conf. Multimodal interfaces - ICMI ’03*, p. 273, 2003.

[44] J. Moll, E.-L. S. Pysander, K. S. Eklundh, and S.-O. Hellstrom, “The Effects of Audio and Haptic Feedback on Collaborative Scanning and Placing,” *Interact. Comput.*, vol. 26, no. 3, 2013.

[45] E. Sallnäs, “Collaboration in multi-modal virtual worlds: comparing touch, text, voice and video,” *Journal of Endodontics,* vol. 28. pp. 172–187, 2002.

[46] C. S. Nam, J. Shu, and D. Chung, “The roles of sensory modalities in collaborative virtual environments (CVEs),” *Comput. Human Behav.*, vol. 24, pp. 1404–1417, 2008.

[47] J. N. Bailenson and N. Yee, “Virtual interpersonal touch: Haptic interaction and copresence in collaborative virtual environments,” *Multimed. Tools Appl.*, vol. 37, pp. 5–14, 2008.

[48] J. Smith and K. MacLean, “Communicating emotion through a haptic link: Design space and methodology,” *Int. J. Hum. Comput. Stud.*, vol. 65, pp. 376–387, 2007.

[49] a. H. Mason and C. L. MacKenzie, “The effects of visual information about self-movement on graspforces when receiving objects in an augmented environment,” *Proc. 10th Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst. HAPTICS 2002*, 2002.

[50] P. M. Fitts, “The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement.pdf,” 1954.

[51] G. A. Miller, “The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information,” *Psychol. Rev.*, vol. 63, no. 1, 1956.

[52] A. E. L. Kaufman, M. W. Lord, T. W. Reese, J. Volkmann, S. The, A. Journal, and N. Oct, “The Discrimination of Visual Number,” vol. 62, no. 4, pp. 498–525, 1949.

[53] I. S. Mackenzie, “Fitts` Law as a Research and Design Tool in Human-Computer Interaction,” *Human-Computer Interact.*, vol. 7, no. 1, pp. 91–139, 1992.

[54] Geomagic, “The Geomagic Touch Haptic Device.” [Online]. Available: http://www.geomagic.com/en/products/phantom-omni/overview. [Accessed: 23-Jun-2015].

[55] M. Poyade, “Motor Skill Training using Virtual Reality and Haptic Interaction: A Case Study in Industrial Maintenance,” 2013.

[56] J. Sauro and J. R. Lewis, *Quantifying the User Experience: Practical Statistics for User Research*. Elsevier, 2012.

[57] “NASA TLX: Task Load Index.” [Online]. Available: http://humansystems.arc.nasa.gov/groups/tlx/. [Accessed: 22-Sep-2015].

[58] S. G. Hart, “NASA-Task Load Index (NASA-TLX); 20 Years Later,” *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 50, no. 9, pp. 904–908, 2006.

[59] D. Wetcher-Hendricks, *Analyzing Quantitative Data: An Introduction for Social Researchers*. John Wiley & Sons, 2014.

[60] H. W. Lilliefores, “On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown,” *J. Am. Stat. Assoc.*, vol. 62, no. 318, pp. 399–402, 1967.

[61] F. E. Grubbs, “Procedures for Detecting Outlying Observations in Samples,” *Technometrics*, Apr. 2012.

[62] J. W. Tukey, “Comparing Individual Means in the Analysis of Variance,” *Biometrics*, vol. 5, no. 2, pp. 99–114, 1949.

[63] D. J. Levitin, K. Maclean, M. Mathews, and L. Chu, “The Perception of Cross-Modal Simultaneity,” *Int. J. Comput. Anticip. Syst.*, pp. 323–329, 2000.

[64] K. E. MacLean, “Haptic Interaction Design for Everyday Interfaces,” *Rev. Hum. Factors Ergon.*, vol. 4, pp. 149–194, 2008.

[65] A. Chan, K. Maclean, and J. McGrenere, “Learning and Identifying Haptic Icons under Workload,” *First Jt. Eurohaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst. World Haptics Conf.*, 2005.

[66] A. Cockburn, B. Mckenzie, and N. Zealand, “Evaluating the Effectiveness of Spatial Memory in 2D and 3D Physical and Virtual Environments,” 2002.

[67] J. A. Langlois, W. Rutland-Brown, and M. M. Wald, “The epidemiology and impact of traumatic brain injury: a brief overview,” *J. Head Trauma Rehabil.*, vol. 21, no. 5, pp. 375–378, 2006.

[68] D. H. Daneshvar, C. J. Nowinski, A. C. McKee, and R. C. Cantu, “The epidemiology of sport-related concussion.,” *Clin. Sports Med.*, vol. 30, no. 1, pp. 1–17, vii, Jan. 2011.

[69] M. S. Dziemianowicz, M. P. Kirschen, and S. L. Galetta, “Sports-related concussion testing,” *Curr. Neurol. Neurosci. Rep.*, vol. 12, no. 5, pp. 547–559, 2012.