Abstract

During extravehicular activities (EVAs), also known as spacewalks, astronauts are exposed to the hazardous conditions of space. Therefore, they must accomplish tasks quickly and have easy access to important information. This study aimed to investigate the effect of heads-up displays (HUDs) on astronaut performance during a maintenance-focused EVA. We first compared users' completion times, comfort, and other factors while they performed operations on a task board using audio instructions, using instructions on an off-the-shelf Microsoft HoloLens HUD, or using a combination of the two. These tests showed a decrease in average mental demand as well as a decrease in mean task completion time for the combined HoloLens and audio as compared to the HoloLens or audio alone. Using these results, we designed and fabricated two versions of a display integrated with an astronaut helmet: (1) a screen system mounted outside the helmet in the lower right of the wearer's comfortable vision range and (2) a projector integrated into the structure of the helmet that projects onto glass in the wearer's upper field of view. By making important task information more accessible, our prototypes have the potential to increase astronaut safety by decreasing the time they spend on EVAs. Results from testing show that users perform better with and prefer a visual display in addition to audio communication. This means a visual display can help reduce the duration of an EVA while keeping the user comfortable and focused.

DEVELOPING AN INTEGRATED HEADS-UP DISPLAY FOR ASTRONAUTS

Team VISOR

Authors:

**Katheryn Fox, Radhika Karsalia, Jillian Kunze,
Christoph Neisess, Zachary Peters, Roshan Rao,
Brady Sack, Matthew Sieh, Ryan Skoletsky,
Shelly Szanto, Matthew Wilkin**

Advisor:

Dr. David Akin

**Thesis submitted in partial fulfillment of the requirements of the
Gemstone Honors Program, University of Maryland, 2020**

ii. Dedication

This thesis is dedicated to Trevor Quinn.

iii. Acknowledgements

Team VISOR would like to thank Dr. Skendall, Dr. Coale, and all of Gemstone staff, Lindsay Inge Carpenter, our librarian, our mentor Dr. Akin, and all of our discussants who reviewed this paper before its presentation. They are Ed Rezac, Lindsay Aitchison, Lemuel Carpenter, Brian Roberts, and Dr. Mary Bowden.

iv. Table of Contents

Abstract
ii. Dedication
iii. Acknowledgements
1 Introduction 2
- 1.1 Motivation 2
- 1.2 Brief Introduction of Heads-Up Displays 2
- 1.3 Research Goals 4

2 Literature Review 5
- 2.1 Spacesuits and Extra Vehicular Activity 5
  - 2.1.1 Current State of Spacesuits and Helmets 5
  - 2.1.2 Description of EVAs 9
- 2.2 HUD Technologies 10
  - 2.2.1 Existing HUD/AR Systems 10
  - 2.2.2 Projector and Screen Technologies for Prototype Development 14
- 2.3 HUDs in Space 18
  - 2.3.1 HUD use during EVAs 18
    - 2.3.1.1 Optics and Ergonomy of HUDs for Astronauts 18
    - 2.3.1.2 Astronaut HUD Experiments 19
    - 2.3.1.3 Power Consumption and Safety 21
  - 2.3.2 How HUDs can Improve Efficiency and Human Experience 22
    - 2.3.2.1 Overview of Human Factors Tests 22
    - 2.3.2.2 Augmented Reality and Procedural Tasks 24
- 2.4 Gaps in Research 27

3 Methodology Results and Discussion of Round 1 and Round 2 Testing 27
- 3.1 Introduction 27
- 3.2 Round 1 Methodology (Preliminary HoloLens Testing) 28
  - 3.2.1 Motivation/Purpose 28
  - 3.2.2 Data Collection Decisions 29
  - 3.2.3 Software HUD Development 29
  - 3.2.4 Round 1 Testing Process 30
- 3.3 Round 1 Results and Discussion 32
- 3.4 Round 2 Methodology (Prototypes/Proof of Concepts Along with HoloLens Testing) 42

3.4.1 Motivation 42
  3.4.2 Hardware HUD Development 43
    3.4.2.1 Version 1 System 43
    3.4.2.2 External Screen Display 47
      3.4.2.2.1 Why use a screen? 47
      3.4.2.2.2 Mounting 47
    3.4.2.3 Helmet-Mounted Display 48
      3.4.2.3.1 Why Use a Projector? 48
      3.4.2.3.2 Projector Selection 49
      3.4.2.3.3 Optical Components 49
      3.4.2.3.4 Design Iterations 50
    3.4.2.4 Secondary Components 55
      3.4.2.4.1 Shoulder Mount Development 55
      3.4.2.4.2 Supporting Electronics 57
  3.4.3 Round 2 Testing Process 57
3.5 Round 2 Results and Discussion 59
  3.5.1 Results 59
    3.5.1.1 Efficiency (Times) 59
    3.5.1.2 NASA-TLX Score Analysis 65
    3.5.1.3 User Preferences (Rankings) 67
    3.5.1.4 Confounding Variables 69
    3.5.1.5 Overall Results Discussion 70
  3.5.2 Discussion (Non-data Observations) 71
4 Conclusions and Future Work 73
4.1 Future Directions for Project 73
4.2 Our Systems and Space Applications 74
5 Works Cited 75
6 Bibliography 80

List of Abbreviations

AHP = Analytical Hierarchy Process
ANOVA = Analysis of Variance
AR = Augmented Reality
CAD = Computer-Aided Design
EMU = Extravehicular Mobility Unit
EVA = Extravehicular Activity
FOV = Field of View
HDMI = High-Definition Multimedia Interface
HMD = Head-Mounted Display
HUD = Heads-Up Display
ISS = International Space Station
IVA = Intra-vehicular Activity
LED = Light-Emitting Diode
NASA = National Aeronautics and Space Administration
NASA-TLX = NASA Task Load Index
OLED = Organic LED
MCH-UVD = Modified Cooper-Harper for Unmanned Vehicle Displays
MDRS = Mars Desert Research Station
RGB = Red, Green, and Blue
USB = Universal Serial Bus
WearSAT = Wearable Situational Awareness Terminal
xEMU = Exploration EMU

1 Introduction

1.1 Motivation

Extravehicular activities (EVAs) are essential for both discovery and astronaut safety. They enable astronauts to directly interact with objects in space, test new technologies, and repair their own vessels and other mechanisms so they can function safely. In 2019, astronauts at the International Space Station spent a total of 145 hours and 36 minutes on spacewalks—the equivalent of about 6 days [1]. The process of preparing and executing a spacewalk is complex, expensive, and hazardous because astronauts are exposed to the harsh conditions of space. It is essential that astronauts have the tools, knowledge, and means of communication that they need to be as time and resource efficient as possible, while maintaining high levels of caution and safety awareness.

As humans look toward the future of space travel, including missions to Mars and beyond, it is essential that astronauts can receive information in a way that does not rely on a direct line of communication with mission control. In the case of a Mars mission, the time delay, which is, on average, about 14 (high of 24, low of 4) minutes one way, will be too great for effective verbal communication of astronaut status and tasks [2]. Space programs are now faced with the challenge of providing astronauts the support they need without constant communication.

1.2 Brief Introduction of Heads-Up Displays

One possible way to give astronauts important information during EVAs is to incorporate a heads-up display (HUD) into astronaut helmets. A HUD is a transparent display that can present data to its user without requiring them to look away from their normal field of vision. This often takes the form of words and images in a digital overlay within some part of the user's field of view. These displays provide the framework of augmented reality (AR) technology, in which 3-

dimensional digital models can be overlaid onto a user's normal field of view. These technologies could benefit astronauts by displaying procedures, diagrams, and descriptions to aid them while they are performing tasks such as maintenance and experiments during EVAs. While astronauts are currently able to complete their EVAs without the assistance of HUD or AR displays, the potential safety and efficacy benefits of the technology continues to motivate research on its deployment in space exploration.

AR technology has already proven useful in education and modeling fields. The Microsoft HoloLens, for example, is a commercially available AR headset which possesses several useful applications in those areas [3]. One such application is HoloHuman, which allows the user to view and manipulate a model of the human body for medical students and researchers to conduct virtual dissections and mockup surgeries. Other applications can create manipulable models of objects or buildings, with uses in fields such as engineering and architecture.

HUDs have never been used in space-based exploration or missions, though several companies have developed their own HUD systems that could be implemented within astronaut helmets. For example, the company Manas has created a heads-up display called Aoudo.X that provides an astronaut with a map of their surrounding area, relaying important weather and navigation information [4]. Several studies have investigated the efficacy of commercial HUDs such as Google Glass or the Microvision goggles in EVA-like scenarios, providing useful guidelines on the requirements for a fully integrated HUD in an astronaut helmet [5][6]. So, while space agencies have not incorporated HUD devices into their missions as of yet, many groups are interested in determining how helpful that technology could be and have begun solving relevant engineering and design problems.

1.3 Research Goals

With digital overlay technologies becoming more advanced and dramatic spacesuit redesigns currently underway such as NASA’s Exploration EMU (xEMU) project, there exists a space for innovation that could lead to significant improvements in astronaut experiences during EVAs. Based on the potential ability of these sorts of devices to alleviate problems during space missions, the overarching mission of our research was to investigate how a HUD could be integrated with modern spacesuits and determining the extent to which this change might make EVAs safer, less costly, and more efficient.

Our study consisted of two broad phases. The first addressed whether the use of a HUD in EVA-style procedures would cause notable improvements in task completion. This phase was a proof of concept to show HUDs were helpful with EVA tasks. We conducted human factors tests on participants who performed sets of tasks that simulated basic maintenance operations; the results could provide a baseline of comparison for later prototypes. We focused on testing metrics that could be important for EVAs such as task completion time, mental and physical demand, number of errors made, and industry-standard survey results. To isolate the effect of our variable (HUD assistance), these tests were conducted by communicating instructions to participants via audio transmission (the control), a commercial HUD, and a combination of the two. In this preliminary stage of the project, we gained experience performing human factors tests and learned how HUD assistance could best help performance on EVA-like tasks.

Based on the results from our initial testing stage, our second phase in our project consisted of developing original hardware and software for prototype HUD systems integrated with spacesuits to help with EVAs. We considered aspects of design such as how information should be presented to subjects for optimal usefulness and how to minimize the intrusiveness of any additions. Based on these requirements, we settled on two independent designs: a screen display external to an astronaut helmet and a projector-based HUD within the helmet. Once hardware and software were completed and integrated for both designs, we compared the performance of our prototypes against the systems used in phase one. Similar human factors tests were conducted to compare performance with the same metrics by isolating each form of communication assisting the subjects in completing the simulation tasks. Using the results from these final tests, we found which designs were most helpful and why, which has implications for the best ways for astronauts to conduct EVAs in the future.

2 Literature Review

2.1 Spacesuits and Extra Vehicular Activity

2.1.1 Current State of Spacesuits and Helmets

Spacesuits provide critical support for astronauts as they explore space. Since their first use in 1959, they have been adapted for various purposes as space exploration has grown to encompass a need for a suit within a vessel, in microgravity, on the Moon's surface, and eventually on Mars.

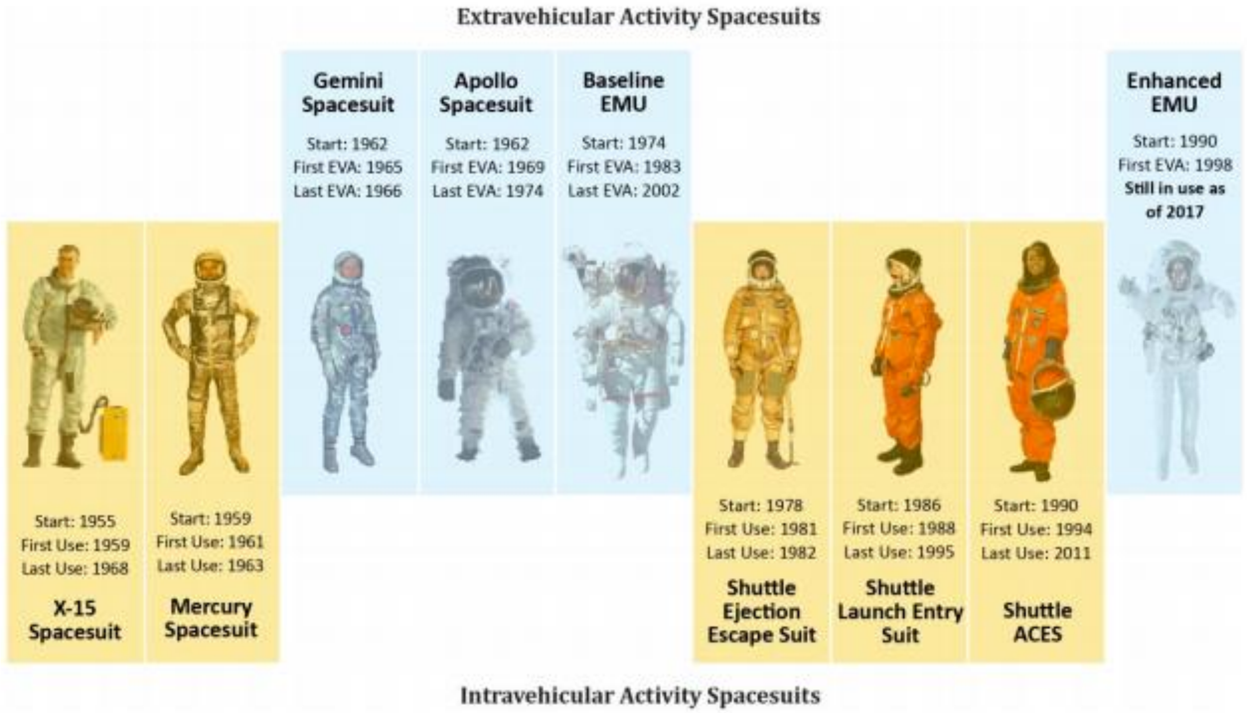
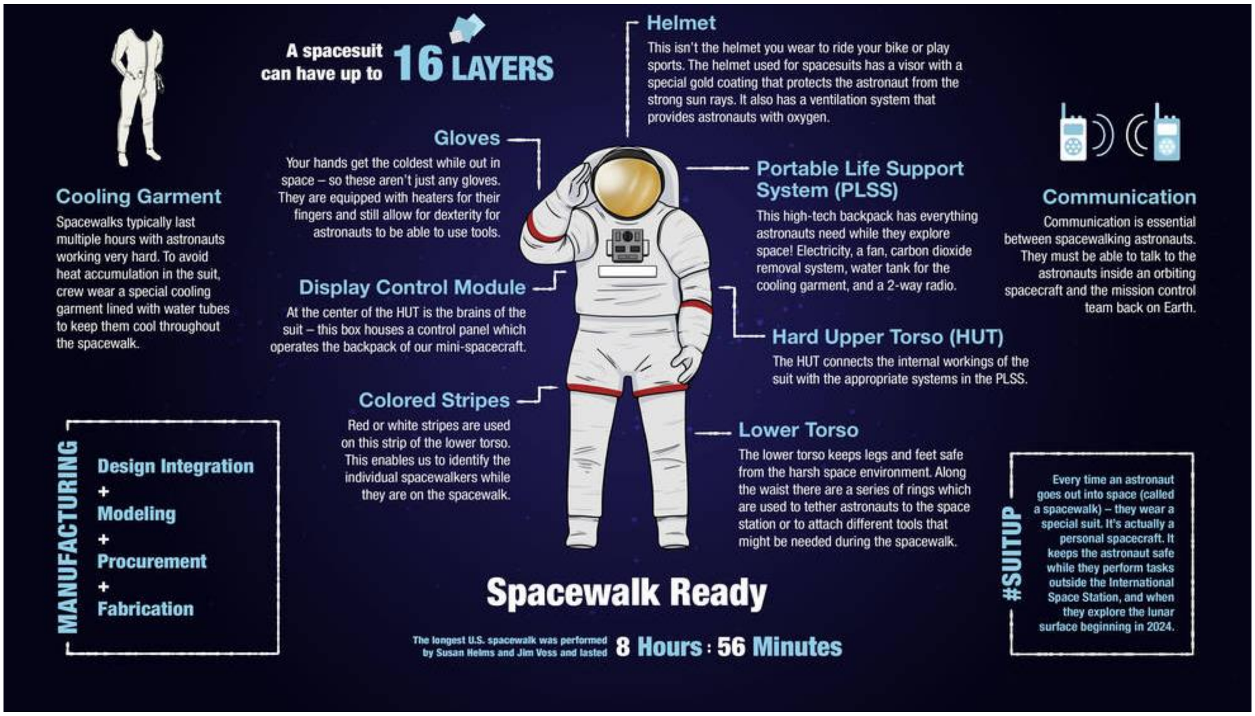


Figure 1: NASA spacesuits from 1955 to 2017 [7].

The spacesuits used today on the International Space Station (ISS) for EVAs are called Extravehicular Mobility Units, or EMUs. Created in 1974 by ILC Dover, the “baseline” EMU was built specifically for microgravity EVAs and provided pressure, thermal, and micrometeoroid protection [8]. It was used by crew members for seven-hour-long EVAs frequently without any negative effects.

With construction beginning on the ISS in 1998, the need arose for a suit that could handle substantially more frequent EVAs [8]. The current design was enhanced from the 1974 “baseline” suit. It was redesigned to include features that made it able to handle more frequent missions and last longer without needing substantial ground maintenance.



Cooling Garment

Spacewalks typically last multiple hours with astronauts working very hard. To avoid heat accumulation in the suit, crew wear a special cooling garment lined with water tubes to keep them cool throughout the spacewalk.

Gloves

Your hands get the coldest while out in space – so these aren't just any gloves. They are equipped with heaters for their fingers and still allow for dexterity for astronauts to be able to use tools.

Display Control Module

At the center of the HUT is the brains of the suit – this box houses a control panel which operates the backpack of our mini-spacecraft.

Colored Stripes

Red or white stripes are used on this strip of the lower torso. This enables us to identify the individual spacewalkers while they are on the spacewalk.

Helmet

This isn't the helmet you wear to ride your bike or play sports. The helmet used for spacesuits has a visor with a special gold coating that protects the astronaut from the strong sun rays. It also has a ventilation system that provides astronauts with oxygen.

Portable Life Support System (PLSS)

This high-tech backpack has everything astronauts need while they explore space! Electricity, a fan, carbon dioxide removal system, water tank for the cooling garment, and a 2-way radio.

Communication

Communication is essential between spacewalking astronauts. They must be able to talk to the astronauts inside an orbiting spacecraft and the mission control team back on Earth.

Hard Upper Torso (HUT)

The HUT connects the internal workings of the suit with the appropriate systems in the PLSS.

Lower Torso

The lower torso keeps legs and feet safe from the harsh space environment. Along the waist there are a series of rings which are used to tether astronauts to the space station or to attach different tools that might be needed during the spacewalk.

The longest U.S. spacewalk was performed by Susan Helms and Jim Voss and lasted 8 Hours : 56 Minutes

Every time an astronaut goes out into space (called a spacewalk) – they wear a special suit. It's actually a personal spacecraft. It keeps the astronaut safe while they perform tasks outside the International Space Station, and when they explore the lunar surface beginning in 2024.

Figure 2: Labelled extravehicular mobility unit [9].

These suits have not seen a drastic shift in design in decades, as they were first designed more than 40 years ago with an expected lifespan of 15 years [8]. One of the key features of the EMU is the Primary Life Support System (PLSS) (shown in **Figure 2**), a backpack-like structure which performs a variety of critical functions while an astronaut is performing tasks on a spacewalk. These PLSSs have far outlasted their planned lifespans, with 11 of the original 18 still in use today on the ISS. While this has been better than expected, it has also raised concerns about astronaut safety as well as potential problems with maintaining inventory as more suits may need to be retired.

Over the past decade, the National Aeronautics and Space Administration (NASA) has spent over $200 million on three new spacesuit development projects, with plans to test them on the ISS before 2024 [8]. The three are the Constellation Space Suit System (IVA and EVA suit) which received $135.6 million, the Advanced Space Suit Project (EVA suit) which received $51.6

million, and the Orion Crew Survival System (launch and entry suit only) which received $12 million. However, due to funding problems and a lack of a clear plan, there have been some doubts from NASA employees, including NASA Inspector General Paul Martin and Space Operations Director Ridge Bowman, surrounding whether or not a complete next-generation prototype will be ready for ISS testing by 2024 [8].

As of 2019, NASA has put out a call to the industry to look for more input on its recent xEMU (Exploration EMU) design [10]. There are many updates being made to the proposed suit, including modularization to fit different body types, the ability to add and remove layers for different environments and temperatures, and a new communication system. Interestingly, though they have shifted away from the "snoopy caps" worn today, they have not yet embraced HUD technology, and have instead replaced the current communication headset with "multiple, embedded, voice-activated microphones inside the upper torso" [10]. **Figure 3** shows the currently published helmet design with sun shields partially deployed.

Figure 3: xEMU helmet design, August 2019 [11].

2.1.2 Description of EVAs

EVAs are essential both for safety and for achieving the mission of understanding and exploring space. Any activity performed outside of one's spacecraft or habitat is considered an EVA. This includes activities in microgravity, such as those performed on the ISS, as well as lunar surface explorations, known as moonwalks. EVAs became very important to the space program during the Space Shuttle era (1981-2011) [12]. During the program, spacewalks became routine because they were essential for construction and maintenance on the ISS.

Since beginning construction of the ISS in 1998, there have been 227 spacewalks on the ISS alone [1]. Some examples of tasks performed on these EVAs include replacing parts, installing new parts, and repairing various objects. Common missions include replacing batteries and repairing scientific tools such as the Alpha Magnetic Spectrometer, which saw four separate repair missions in 2019 and 2020. These missions require intense planning and preparation from everyone involved. For astronauts, they must have fine motor control of various instruments and small parts such as hand drills and bolts, as well as a precise understanding of what they must do and when, all while remaining vigilant of their suit and surroundings to ensure their own safety.

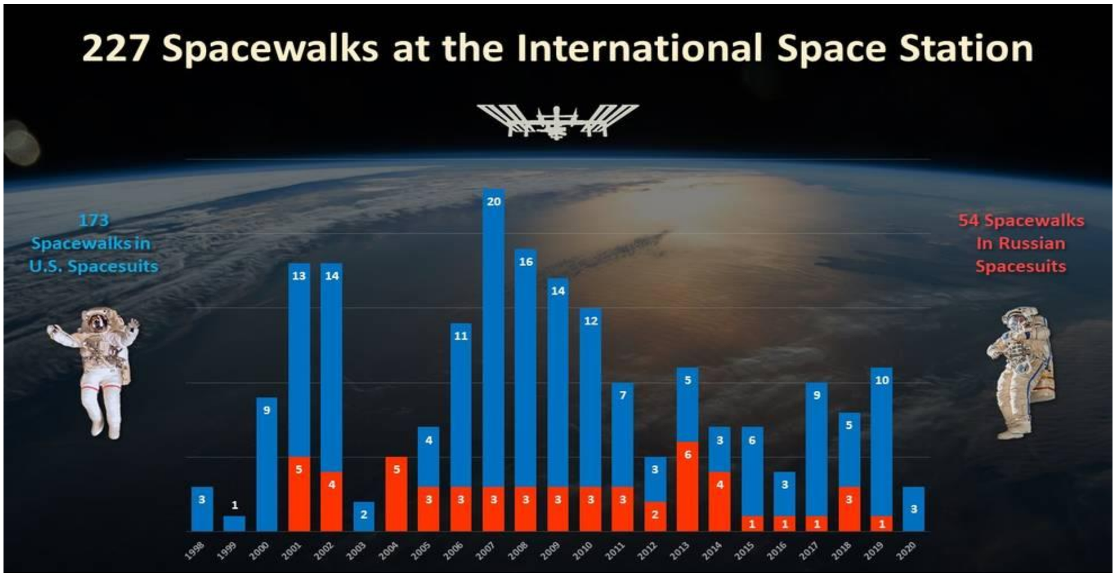


Figure 4: Graph of spacewalks at the ISS (1998-2020) [1].

These missions are hazardous and expensive. It is impossible to predict every possible problem that could arise on any given mission. Human error, suit malfunctions, and equipment failure are all possible every time an astronaut exits the security of the spacecraft. Even recently, during a spacewalk on January 15, 2020 one of the astronauts' helmet lights and camera became unattached [13]. Unable to reattach them, her fellow astronaut clipped them to her tool bag. They were able to continue and complete the mission goals, though ground control cautioned them to be careful, telling her she was "missing that additional protection." Another example of an unanticipated situation occurred during the spacewalk on January 25, 2020, where the astronauts were delayed due to a strap stuck between the crew lock section and the equipment lock area, creating an air leak [14]. This prevented the depressurization of the airlock, and caused them to restart the process. While not life-threatening, these examples show that unforeseen circumstances are not only possible, but happen quite frequently and demonstrate some of the problems possible on EVAs. It is essential that astronauts have at their disposal every possible tool to ensure their safety and wellbeing, particularly as we look towards redesigning spacesuits for the next generation of astronauts.

2.2 HUD Technologies

2.2.1 Existing HUD/AR Systems

A HUD is defined as any transparent display that presents data to the user without requiring them to look away from their normal field of vision. Eliminating the need to look away from the workstation, providing full use of both hands for work, and increasing the amount and types of data able to be displayed, a HUD-integrated helmet visor would increase the efficiency and safety of astronauts during EVAs. A HUD also lays the framework for the future integration of AR, a field in which NASA has taken great interest because it could benefit astronauts by displaying three-dimensional visualizations of procedures, diagrams, and descriptions while they operate and repair equipment. While astronauts have been successful thus far without the assistance of a HUD or AR, the potential benefits of the technology justify its research and development.

HUDs have been a subject of extensive research in recent years. The first HUDs were helmet-mounted displays that were developed in the 1970s for pilots' helmets as shown in **Figure 5** [15]. The goal of transparent HUDs is to provide complementary information while supporting the wearer's focus towards action and/or increase situational awareness in order to provide additional functionality to human senses. Furthermore, a HUD with head or eye trackers could be used for AR purposes, which provides a higher level of understanding and information context [15].

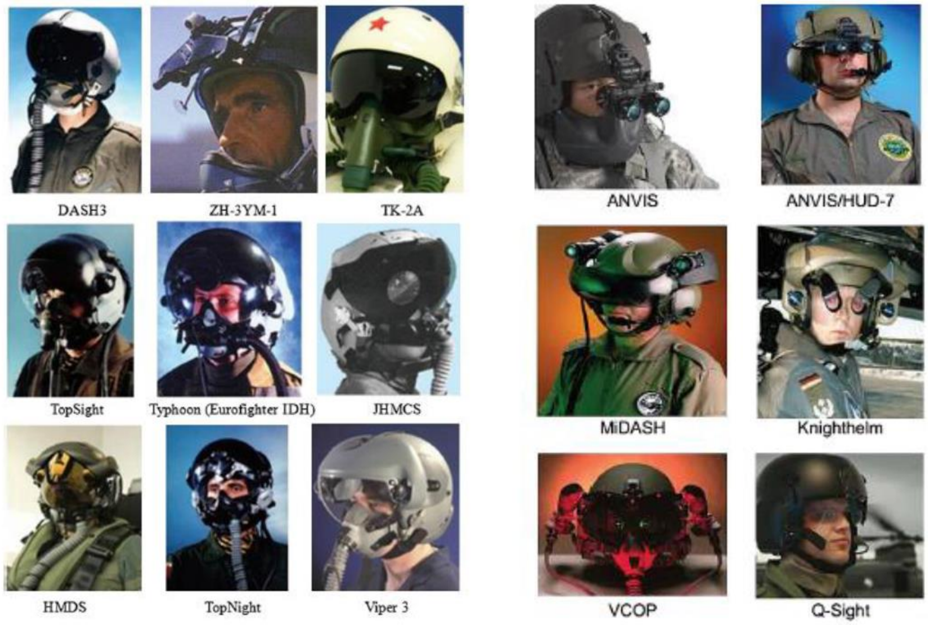


Figure 5: Different types of fighter pilot HUDs [15].

Researchers have developed many HUD interfaces specifically for navigation cues while on the road such as navigation arrows, freeway exits, points of interest, and hazard alerts, utilizing anything from small areas to the entirety of the windshield [16]. The commercial sector has taken an interest in this opportunity as well. One example of a commercial HUD is the Navdy, a portable device with audio and gesture input that includes navigation, audio, and cellular features [17]. Additionally, helmets and screens have been developed to aid motorcycle drivers and industrial workers [17]. Researchers from Keio University have developed a HUD interface using a laser pico-projector and acrylic board to learn the optimal position for a display and amount of information to provide drivers [18]. This and other studies have been used in the corporate sector to develop helmets such as the Skully motorcycle helmet, which has blind spot cameras and calling features, and the Daqri helmet, which is used in industrial settings to aid in maintenance, repair, and inventory [16]. Many similar technologies are being developed for human use in a variety of occupations, such as modern fighter pilots, divers, fire fighters, and national defense officials [5].

One specific and well-known device is Microsoft's HoloLens. The Microsoft HoloLens is a projector-based AR system that displays a high-resolution image close to the user's eye on an otherwise transmissive screen [3]. The system Microsoft has developed to do this is incredibly compact, and does not require extensive support electronics or other physical constraints such as a high power draw or high heat generation. This means the HoloLens can operate independently of a cable connection and is light enough to be worn on a person's head without discomfort. The HoloLens also makes use of a dedicated secondary graphics unit to take the load off of the primary central processing unit, so the latter can focus more on functioning like a regular computer and less on driving the projectors. Though the materials for a basic understanding of the HoloLens' function and layout are publicly available, Microsoft has not chosen to release many of the technical specifications. However, Microsoft's publications do show that the basic function of the HoloLens is similar to the system we aim to develop in that it is able to display information in a customizable manner.

Therefore, we determined that the HoloLens, shown below in **Figure 6**, was the ideal benchmark from the available consumer-accessible systems for initial testing. For example, the HoloLens is currently the only commercially sold AR device with a transparent screen, and having good external visibility was one of the basic requirements for our prototypes [3]. The HoloLens is also small enough to fit into a modified spacesuit helmet, meaning we could perform our initial testing inside a helmet. These assets permitted us to quickly design and test a simple HUD using Unity, a cross-platform game engine which can be run by the HoloLens.

Figure 6: The Microsoft HoloLens [3].

Despite the HoloLens technical details being proprietary, there are several articles documenting experiments and studies conducted using the HoloLens. Among these are studies that compare the effectiveness of using the HoloLens versus other systems to complete a list of tasks [20] and testing the real-time resolution of projections created using the HoloLens [21]. Many of these studies have used the HoloLens to cover topics similar to ours. One relevant experiment was *Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks*, in which researchers created a series of tasks using Lego assembly instructions for participants to complete. Participants were asked to write down comments and observations on the hardware and task process, similar to our preliminary testing with the Hololens [20]. This furthered our confidence in the HoloLens's ability to provide a solid baseline on which to base our prototype design and development.

2.2.2 Projector and Screen Technologies for Prototype Development

The technical considerations of EVAs inform and guide the design process of spacesuits, setting engineering requirements that any helmet visor must meet. Working within those requirements, we considered and eventually employed technology from two promising theaters of HUD development: projection and screens. Projection uses a separate device to display images on a surface, while for a screen system the screen itself creates the display. Multiple manufacturing processes exist for both HUD types, each with their own benefits and drawbacks in the context of a spacesuit helmet for astronaut use.

A projection-based HUD system has two major components: a projector and a reflective or semi-reflective surface. Some small projectors make use of three light-emitting diodes, or LEDs, in red, green, and blue (RGB) as a light source. The light from these LEDs is then collected and integrated before being emitted through the lens of the projector. LED-based projectors are simple in that they require less light filtering compared to older units that use mercury arc discharge lamps. They also provide the benefit of controlling the ratio of RGB light emitted by regulating the current to each LED. Furthermore, LEDs do not consume much power, and have a relatively long lifespan compared to conventional lamps. However, a disadvantage of LED projectors is that they do not emit light at the same intensity as some other projectors [22]. There are generally many options for screens to project onto, but within the context of HUDs it is necessary to have a partially, or fully, transparent screen.

A recent development in the projector field is retro-reflective projection technology. A retro-reflective surface consists of beads about 50 micrometers across that each bounce back light in the direction from which it came, as opposed to scattering it or reflecting it at an angle. Retro-reflective surfaces can be applied to a transparent object, meaning the resulting screen will appear transparent when there is no image projected onto it [23]. This method of reflecting light allows the projectors to be placed to the side of the user's eyes and still effectively display images; thus, the projectors would not have to occupy too much space inside an astronaut helmet.

Aside from the projection screen and the projector itself, the most important component in a projection-based setup is the mounting system. This must be designed to be secure, yet flexible enough to allow the device to move, especially for a mounting system in a helmet as the helmet may undergo substantial movement or small impacts. The goal of the mount is to keep the projector and the surface it projects onto in optical alignment so the resulting image does not get distorted. This is often done through brackets that only allow for motion in certain directions, and which resist twisting or flexing [24].

Alternatively, screen technology provides a second possible technique for developing HUDs. This could be implemented as a transparent screen incorporated into a helmet visor or an opaque screen mounted to the exterior of a helmet. Most methods to create screen technologies involve layering a glass-like screen with various chemicals, elements, or materials [25]. This technology differs from projection in that the informational display is generated from within the visor itself, rather than from another device.

One promising screen technology is organic LEDs (OLEDs), which are a subset of LEDs that have an electroluminescent layer composed of an organic compound. While inorganic LEDs are backlit by an additional luminescent layer, OLEDs contain sufficient organic electroluminescence to produce a visible display without a backlight, reducing power consumption, mass, and complexity [26]. Furthermore, OLEDs can be fabricated on plastic and flexible substrates, which permits the creation of curved and transparent screens [27]. Due to their flexible nature, these displays also have substantial resistance to shattering. Because of the isolated and energy-intensive nature of space exploration, any proposed technology for this field must adhere to certain mass and power limitations; OLED technology would likely meet those requirements. However, OLEDs are currently not very transparent, and therefore would not apply very well to HUDs. Future research may elevate this technology to make it more suitable for HUDs, but the current state of OLEDs was not useful for our prototyping.

Resonant nanoparticle scattering technology is another principal type of transparent screens that shows significant promise. Nanoparticles are microscopic pieces of matter with sizes on the scale of nanometers (usually, smaller than 100 nanometers). By exploiting a natural property called resonance of certain types of nanoparticles, a completely transparent display can be produced for various applications. Characteristics of nanoparticles cause them to selectively scatter, or redirect, light that lies within an extremely small range of wavelengths surrounding what is called the nanoparticles' resonant wavelengths [28]. Each type of nanoparticle has its own unique resonant wavelengths, analogous in many ways to the spectral lines of elements and compounds.

Particular nanoparticles can be selected based on their resonant wavelengths and embedded into a transparent medium via nano-scale construction. Because the nanoparticles do not influence light aside from their resonant wavelengths, and since the medium is already transparent, ambient light moves through with statistically insignificant scattering or absorption, making the medium see-through. However, light at the resonant wavelength(s) scatters noticeably, an effect consequential enough to make it visible on the medium [28]. With such behavior in mind, it is a natural next step to consider whether images could be displayed on a transparent medium. A 2014 study conducted by Wei Hsu and his research team investigated a transparent medium consisting of a polymer matrix with embedded silver nanoparticles [25]. They found that any chosen image could be shown on the medium as long as the picture was monochromatic; in this example, the resonant wavelength corresponded to blue light.

Helmet-mounted projector systems have several advantages for use in a HUD. The display can be projected onto a single, fixed area of the helmet visor; if the projector is small enough, it would not take up too much space inside the helmet. It may be possible for the projectors to be hooked up to an external power source similar to those used for current headlamps and status indicators on the suit, but the wiring involved could also be a safety risk. For sake of cost, weight, and available space, it is most desirable to have the lowest amount of equipment possible in the system without sacrificing too much in quality of the image. Based on all this, we determined that an interior helmet-mounted projector approach could certainly be feasible given the right optical path and hardware design.

After examining the various options for transparent screens, including the promising resonant nanoparticle scattering just described, we decided that approach was not feasible for our research at this time. Although there does exist technology that could make transparent-screen HUDs possible (and which a technology agency like NASA could investigate with far more resources), based on the team's limitations in budget and time, we chose to discard the possibility of employing an integrated transparent screen, though we did use a non-transparent screen in an external screen display setup. Overall, research has shown that HUD technology is advanced and accessible enough for experimentation concerning its application for EVAs, and we found two possible methods of creating a display to assist in tasks: interior helmet-mounted projection and an externally mounted screen. These approaches were used to create our prototypes for testing, as will be explained in the following sections.

2.3 HUDs in Space

2.3.1 HUD use during EVAs

Though HUDs have never been used during an actual EVA by astronauts in space, researchers have investigated the possibility of astronauts using HUDs via Earth-based experiments. Creating an ergonomic HUD system within the confines of an astronaut helmet is a challenge different from the development of less specialized HUDs. The existing field of view within the helmet acts as a design constraint, as do safety concerns. Experiments using commercial HUDs for simulated EVA tasks have identified additional special areas of concern, including glare and adjustability in HUD sizing.

2.3.1.1 Optics and Ergonomy of HUDs for Astronauts

The optics and ergonomy of a HUD in an astronaut helmet is an essential consideration, as astronauts must be comfortable and have an unobstructed field of view to effectively perform their tasks. The operational requirements imposed by the different activities astronauts do while in the unique gravitational environment of space impose unique challenges for astronaut HUD designs. The NASA Human Integration Design Handbook provides specific standards for the visual performance of helmets, including that the helmet's visor must not interfere with the astronaut's ability to accomplish their tasks or present a safety hazard [29].

One major area of consideration is the astronaut's field of view (FOV), as any informational display must not obstruct the astronaut's FOV or subject it to fogging from the wearer's exhalations [29]. An additional important aspect of a display is its stability. For projector systems, the display must be completely stable in order for the projected information to remain centered. Otherwise, there is a risk of part of the display being cut out of the image given to the astronaut, potentially causing the loss of critical information. If the information is presented on an external screen, the system should not be dislodged by accidental impacts or the astronaut's own movement.

In a HUD, the apparent depth is the distance from the user's eyes at which the displayed information appears to sit, also known as the focal point [5]. Displays can become uncomfortable when they do not follow the user's natural focus cues, forcing the user to frequently readjust their focus when looking back and forth at the display and the surrounding environment. This phenomenon, called Vergence-Accommodation Conflict, can cause visual discomfort, eyestrain, nausea, headaches, and even pathologies in children's developing visual systems. To reduce this impact of this issue, some technologies now incorporate a focus-tunable mode that gives the user an adjustable focal point for the display.

2.3.1.2 Astronaut HUD Experiments

The Mars Desert Research Station (MDRS) in Utah published a study entitled *Ergonomy of Head Mounted Displays Inside Analog Spacesuit - Mars Analog Extravehicular Activities* [15]. The authors of this study saw the goal of transparent Head-Mounted Displays (HMDs) as providing complementary information that enhances the functionality of the user's senses, supporting their focus and increasing their situational awareness. One of the several commercial HUD systems investigated in the study was Google Glass, which projects information onto a crystal placed above the direct sightline of the right eye. Another was a HUD called 4iii, which is a wireless system with a display placed under either of the wearer's eyes. To test the efficacy of the Google Glass and 4iii systems in an EVA-like scenario, the authors had test participants wear the HMDs inside a simulated spacesuit helmet while following a short path and collecting data needed to map the terrain. The two HMD systems were worn in tandem and used for navigation and heart rate monitoring.

To analyze the HUDs efficiency in the simulated EVA, the researchers took videos of the test participants and collected data on head position, Google Glass activity, glare, and other parameters [15]. A complaint from the study participants was that it was difficult to focus on the Google Glass display in direct sunlight. As astronauts must be able to work in all levels of light, and the level of sunlight can change dramatically throughout the orbit of a satellite, this posed a major issue for incorporating Google Glass-like hardware into an astronaut helmet. The authors proposed that placing a transparent foil on the helmet or Plexiglas in the horizontal plane of the Google Glass would be sufficient to provide contrast from the Sun’s glare. Other issues with the HUDs used during this experiment included the lack of a mechanism for personalized adjustment and problems with voice commands due to poor acoustics in the helmet. In addition, the MDRS study revealed that the Google Glass did not work well for all head sizes and, when worn in the helmet, had an insufficient field of vision in the horizontal plane.

Two participants in the MDRS study were asked to rate their physical and mental comfort while wearing the HUD system, as well as other qualitative factors [15]. Based on feedback that the system was difficult to put on, the researchers recommended that future HUD systems for astronauts include an adjustable head piece that could be stretched or narrowed for a better fit. Users also noted that the display required dedicated focus, such that the wearer had to stop operations to focus on the display. Other areas of concern included an imbalance in the display, overheating, poor focus within the field of view, and imprecise responses from the devices to verbal commands.

In the study *Results of a helmet-mounted display precursor demonstration at Desert RATS*, researchers compared the effectiveness of two different HUD devices at assisting astronauts with tasks [30]. The devices were a monocular "see-through waveguide" Microvision goggle, which overlaid images on the view outside the helmet, and an occluded view goggle from Recon Instruments. Though the study did not attempt to integrate these devices into an astronaut helmet, it tested their attributes that may be useful for space applications. In one test, a detailed procedure that astronauts might perform was simulated using the Lunar Science Experiment Package mockup. Researchers found that both the Microvision and the Recon Instruments goggles helped guide the users through tasks on the mockup by displaying diagrams and procedures. Color coding was an effective way to convey information's importance, with green for nominal info, yellow for caution, and red for warning.

The study, *A Wearable Computer for Support of Astronaut Extravehicular Activity*, developed a prototype of a "wearable situational awareness terminal," or WearSAT [6]. The WearSAT system, which runs on a wireless network, can display information and graphics at an eye level during spacewalks using a near-eye micro display that converts a Video Graphics Array (VGA) signal into a 320x240 display on a pair of glasses. The researchers concluded that voice-based communication could be replaced with WearSAT's display, though research was still needed to make its software compatible with the network used on the ISS. These previous astronaut HUD studies demonstrated that incorporating a HUD display with astronaut tasks was feasible, and provided useful results on design constraints that could improve the experience of test participants.

2.3.1.3 Power Consumption and Safety

When designing electronic components for spacesuits, it is essential to ensure they can be installed within the framework of existing suits without extensive modifications. One important aspect of novel electronic components is their power consumption, as the power supply in spacesuits must be able to last the full length of a spacewalk (up to eight hours) [31]. Battery power must be distributed between all of the electric systems in the suit, including the life support system, cooling fans, and the communications assembly. To avoid significant reconfigurations in the battery or power distribution systems, electronic additions to the suit such as a HUD must not draw extensive amounts of power. While our prototype development was not within a full spacesuit and so had less limitations on power consumption, it is still useful to consider how concerns of power draw could affect future iterations of HUDs for astronauts.

The NASA Human Integration Design Handbook gives descriptions of safety hazards to avoid when designing hardware to be used during EVAs [29]. Any surface that an astronaut's bare skin may be exposed to must not exceed 44 degrees Celsius, so as not to go above the human tolerance for heat pain. As a HUD system may have components inside an astronaut's helmet that have the potential to contact their skin, it must follow this restriction. In addition, the voltage pull of the HUD system must not be so great as to pose a risk for electrical shock to the astronaut; NASA sets the maximum exposure voltage in astronaut applications at 32 volts root-mean-square [29].

2.3.2 How HUDs can Improve Efficiency and Human Experience

2.3.2.1 Overview of Human Factors Tests

The NASA Task Load Index (NASA-TLX), developed at the NASA Ames Research Center by the Human Performance Group, is a measure of assessing workload [32]. NASA-TLX rates workload on six subscales: Physical Demand, Mental Demand, Temporal Demand, Effort, Own Performance, and Frustration. Participant's ratings on the subscales are weighted according to how the participants rated the importance of that scale for a particular task. The weight and the rating a participant gives a task may or may not exhibit covariance; for example, participants can say a subscale is very important for a certain task, but also say that the demand on that scale is low.

The Cooper Harper scale was originally developed in the 1960’s to measure the handling qualities of piloted aircraft [33]. It was a way to translate qualitative pilot impressions into quantitative assessments that could be compared and analyzed. Pilots were able to evaluate aircraft by following along a flowchart, which branched based on their experiences with handling qualities such as workload, controllability, and demands [34]. The flowchart ends at a value from 1 to 10, with 1 indicating excellent quality and 10 indicating a system with major deficiencies.

Modified Cooper Harper scales have been developed for systems other than piloted aircraft, such as unmanned vehicle displays [34]. The Modified Cooper-Harper for Unmanned Vehicle Displays (MCH-UVD) scale, shown in **Figure 7,** was developed in 2006 to measure information acquisition, information analysis, and decision making by the operators of these vehicles. It is also arranged in a flow chart so that it could be presented to operators of unmanned vehicles after a test to quickly assess their experience. The MCH-UVD flowchart has been through several iterations, with improvements being made to reduce overly technical language and removing overlap between different flowchart options.

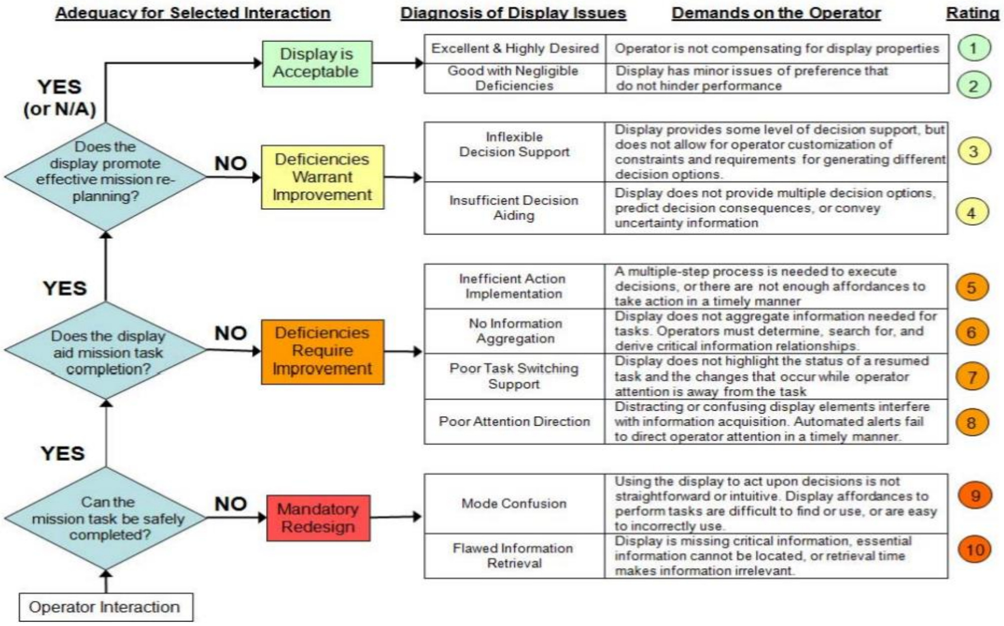


Figure 7: Modified Cooper-Harper chat for unmanned vehicle displays [34].

In the MCH-UVD scale, displays rated 10 or 9 did not provide necessary information or interfered with operations [34]. Displays rated 8, 7, or 6 made it difficult for operators to analyze the information presented due to distracting elements, lack of support for moving between tasks, or incomplete information aggregation. A rating of 5 corresponded to a display with inefficient processes for taking actions. Ratings of 4 or 3 indicated a deficiency in resources needed for decision-making; for example, the display might not predict outcomes or allow for customization to narrow down decision options. Finally, displays rated 2 or 1 were acceptable for information acquisition, information analysis, and decision making and were therefore either good or excellent displays based on the criteria of the MCH-UVD scale.

2.3.2.2 Augmented Reality and Procedural Tasks

In a recent study at NASA's Jet Propulsion Laboratory, researchers investigated how AR instruction would affect completion time of tasks, workload, and number of errors in spaceflight-analogous procedural work [35]. Procedural work is defined as performing a well-defined sequence of activities to accomplish a certain task; this is common in maintenance tasks that one might accomplish in a space exploration scenario. Using AR could potentially mitigate some of the cognitive load of reading and interpreting procedural instructions by superimposing the instructions onto the environment. This could save time by reducing the attention shifts of the user and alert the user to important physical objects by superimposing a box or label on them.

Researchers used the Microsoft HoloLens to display AR instructions for installing and maintaining a mockup of a science instrument used on the ISS [35]. The 20 test participants performed tasks twice, once with the HoloLens and once with paper instructions, attempting to be as fast and accurate as possible. The researchers timed each trial with a stopwatch and counted the number of errors. After each trial, participants completed the NASA-TLX and the System Usability Scale to rate the usability and workload of the instructional method, as well as filling out questionnaires about the pros and cons of each method and stating their preferred method.

The completion time of each trial was significantly faster the first time the participant used the AR instructions than the first time they used the paper instructions [35]. However, for the second time users, completed tasks had no such time difference, as using an AR procedure first was able to train the user to complete the task more quickly with paper instructions. The users committed few errors overall, and there was not a significant difference in errors between the two instructional methods. The mental workload of the task was significantly smaller for the AR instructions, though both instructional methods were rated the same for usability. Participants did have several complaints about the HoloLens: it was heavy, its field of view was small, and they had to strain to shift focus between the physical device and the AR display. But overall, the study suggested that spaceflight operations could improve via the use of AR display of procedural information.

A separate study compared the use of AR instructions versus paper-based instructions for a standardized assembly task [20]. The researchers used a manual assembly task with Lego Duplo bricks to compare AR instructions on a HoloLens, AR visualizations on a smartphone, two-dimensional in-view pictorial instructions on an Epson Moverio HMD, and instructions on paper. While test participants completed the Lego Duplo tasks, researchers recorded the number of mistakes they made, and their task completion times. The 24 participants also filled out a NASA-TLX for each of the four instructional systems.

The study found that on average, participants using the paper instructions took the least time to locate and put their hand in the bin with the correct Legos for their current step [20]. The time needed to pick out the correct brick was similar for all the methods, though paper instructions were once again the fastest. Paper was also fastest for finding the correct position of a brick and assembling the current step. Participants using the smartphone display were the slowest for all these timing phases. The relative amount of errors made with each system differed with different tasks; for example, the fewest errors were made with the HoloLens when finding the correct brick container, but the fewest mistakes were made with paper instructions when assembling the brick of the current step.

Participants in the study perceived the lowest amount of cognitive load as measured on the NASA-TLX using paper instruction, followed by the Epson Moverio HMD, the HoloLens, and the smartphone [20]. The smartphone was criticized for being annoying to handle and difficult to interpret. Half of the users commented that the HoloLens’ field of view was too small, but some also remarked that its three-dimensional display felt natural and easy to use. Participants found that the Epson Moverio display interfered with their visibility and in some cases gave them headaches. Based on their results, the researchers proposed that an AR display like the HoloLens could be used in combination with pictorial instructions, as each instructional method was less error-prone for different tasks.

2.4 Gaps in Research

A large portion of the existing HUD research involves test subjects completing trials wearing a HMD without a spacesuit helmet. The helmet introduces additional constraints on the user and will likely affect task completion time and NASA-TLX responses. Furthermore, many studies do not include audio instructions as a control, which is the current method for communication with ISS astronauts on EVA. There is therefore a need to further explore how displays mounted in spacesuit helmets compare to current communication methods.

Given the review of current literature, we decided to focus on the viability of a HUD inside a spacesuit helmet. The goal of the research described below is to justify the implementation of an AR system in spacesuits and their use on future exploration missions. Due to financial and time constraints, our team decided it infeasible to develop a HUD capable of being used in an actual spacesuit. Therefore, we set out to determine whether or not a HUD is feasible in space applications from a human factors perspective (that is, whether or not a HUD improves task performance, decreases task time, or decreases the number of errors committed) and, in doing so, to justify the potential benefits of HUDs in space applications to future researchers.

3 Methodology Results and Discussion of Round 1 and Round 2 Testing

3.1 Introduction

To quantitatively and qualitatively measure the effects of a HUD on procedural EVA work, we developed an analog EVA task board (**Figure 8**) and had test subjects run through three trials, during which they were asked to work through a set of instructions delivered each time via a different system configuration: audio only, Microsoft HoloLens only, and HoloLens with audio. The audio only system served as the control, as this is the current method used to deliver instructions to an astronaut when they are completing a novel task. If the test subjects demonstrated a decrease in time needed to complete the task without a significant increase in discomfort, then we could justify the use of HUDs in spacesuit helmets.

The following sections outline in detail the procedures and equipment used to conduct the tests. Further details are provided regarding the task board, the software used to develop on the HoloLens, and the data collected.

3.2 Round 1 Methodology (Preliminary HoloLens Testing)

3.2.1 Motivation/Purpose

Our first round of testing consisted of three different tests: audio only, HUD only, and HUD and audio combined. The HUD in this phase was a software system we designed to run on off-the-shelf hardware. We decided to use this setup because our preliminary testing was mostly focused on fine-tuning our software system, while also learning how a HUD system compared with the system currently used on the ISS (audio only). We also wanted to get a sense of whether or not a combined HUD and audio system would be effective. These tests would set up our second round of testing with our own prototype hardware.

For our preliminary testing, we decided to use a Microsoft HoloLens based on a few different factors. First, the HoloLens met many of the requirements we set for an AR headset; it was user-friendly, it was easy to read, and it was compatible with Unity, the software we would use in later HUD development. Using the HoloLens in this phase allowed us to get a good comparison with the audio only system that is currently used on the ISS for EVAs. Through this first round of testing, we also refined our data analysis techniques to be ready for our second round of testing, as well as practiced testing human subjects on our task board.

3.2.2 Data Collection Decisions

In round 1, we collected several data points during tests with audio only, the HoloLens only, and audio and HoloLens combined. These points included the total time each test took to complete, the number of errors made by the test subject, NASA-TLX scores, Cooper Harper ratings, and a few survey questions including a ranking of the three forms of communication from best to worst. These data points allowed assessment of how the HUD affected and the user's performance.

Each test was treated the same way and, therefore, the display for the two HoloLens tests were kept consistent. The only modification done for each test was the randomization of the tasks themselves. The algorithm of the sequence of tasks was made to randomize each time the display refreshed to help mitigate some of the learning curve experience with doing similar tasks three times in a row. We did still observe some learning.

3.2.3 Software HUD Development

The software for each of the HUDs was developed in Unity, a cross-platform game engine that was chosen for its compatibility with the HoloLens. Unity provided a simple, drag and drop platform for designing the layout of the display and interfacing with it via a C# API. To build for the HoloLens, we used the Mixed Reality Toolkit, an open source toolkit developed by Microsoft [source: https://github.com/TeamVISOR/Software]. The other displays needed no additional libraries and were built as Microsoft Windows apps.

Instructions were grouped into four categories and added to respective array lists: buttons, dials, switches, plugs, and carabiners. At each step, a category (i.e. arraylist) was randomly selected and then an instruction was randomly pulled from the arraylist. After it was displayed, the instruction was deleted from the arraylist. Once an arraylist was empty, it was also deleted. Given the uneven size of the arraylists and the fact that they were equally weighted with regards to their probability of being chosen, this program structure tended to lead to the last 15 instructions or so being only buttons.

The HoloLens app was streamed to the HoloLens using the Holographic Remoting feature in Unity, which often resulted in a shaky display.

3.2.4 Round 1 Testing Process

The preliminary testing consisted of three trials performing the same set of 50 simple mechanical tasks in a random order, each time using a different form of communication to provide instructions: one with audio only, one with the HoloLens only, and one combining both. In addition, for each participant, the order of these trials was also randomized to eliminate as much as possible the role of a learning curve stemming from participants beginning or ending with certain communication forms. The pool of subjects was a random selection of 30 adults (chosen to obtain a sample large enough for statistical significance, but small enough to perform testing in a reasonable amount of time) who were recruited using physical and online advertisements, then screened with a survey before beginning testing. The requirements for individuals to qualify as our test participants were that they must have had either 20/20 or corrected vision, had normal hearing, and were not colorblind. If, on the survey, a potential participant marked that they were not sure of their vision, a standard eye exam and color blindness test were used in order to determine their eligibility. We also asked participants their height so that we would be able to see if taller- or shorter-than-average people might have more or less difficulty with seeing and making the most of our display.

Our testing board is organized to resemble a control panel that might appear on a space station. The board contains two sets of switches, five dials with numbers in increments of one, and twenty-five buttons of five different colors. The layout of the testing board in its early stages of development, with labels for each part, is shown in **Figure 8**. In addition, the testing board has an industrial outlet with a corresponding plug and a single metal box, whose lid has four screws (a screwdriver was provided for the participants) and which has inside it four different carabiners. We decided on these objects to mimic the kind of potential actions that an astronaut could encounter while on an EVA. Finally, the board itself is angled towards the test participants by a support structure made of 80/20 aluminum extrusions.

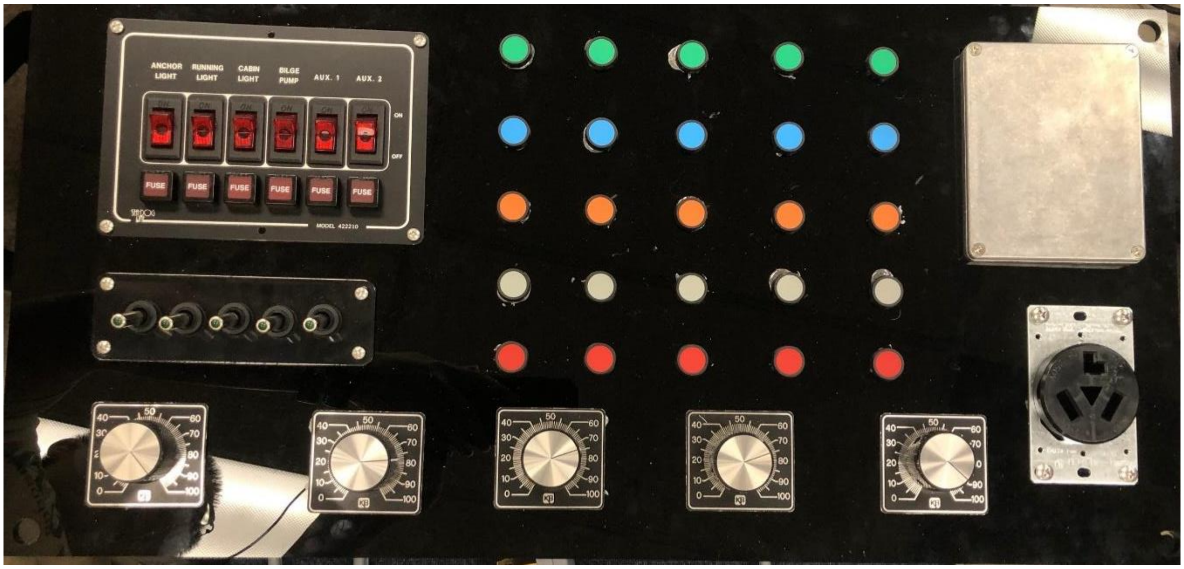


Figure 8: The layout of the "control panel" board which participants used during testing.

In each of the three tests, participants were run through a randomized procedure of 50 steps involving pressing buttons, flipping switches, turning dials to specified numbers, unscrewing screws with a screwdriver, and otherwise interacting with the hardware on the testing board. For example, they were asked to unscrew the lid from the metal box, and then later asked to clip each of the four carabiners found inside onto their designated corners of the testing board before returning them to the box and screwing the lid back on. While the steps to open the metal box and remove the carabiners were deliberately made the very first steps (and last for vice versa), every one of the remaining steps was made completely random to ensure a different procedure for each test.

All participants completed the procedure a total of three times, once for each of the communication methods. The audio communication trial involved direct person-to-person interaction with a specific researcher, as one of the team members would read off the directions to the test subjects. The subjects were able to ask their researcher to read the current step again if necessary. For the HoloLens tests, instructions were displayed without any audio in the subject's field of view as part of the overall AR display, which also incorporated pictures of the parts involved in each task. As an example, the subject might see the directions "turn dial 2 to 75" (for our dial tasks, the participants were instructed to turn each dial to multiples of five to avoid unnecessary precision and difficulty), and the dial would be displayed. The user was expected to use voice commands to advance through the procedure, and was able to view previous steps as necessary. When both visual and audio communications were used, subjects would have access to both the text and images from the HoloLens, as well as direct contact with a team member reading the procedure aloud. The data collected were the completion time, number of errors, original survey questions, and modified NASA-TLX and Cooper-Harper results for each trial. Data were collected for all 30 subjects, and the results will be discussed in the next section.

3.3 Round 1 Results and Discussion

One of the main goals of our testing procedure was to gain a thorough understanding of how different styles of audio and visual communication affect the overall performance of a test subject while performing a series of tasks that would be representative of EVAs. Major portions of the round one data analysis discussed in this thesis can also be found in our 2019 ICES conference paper, *Development of a Heads-Up Display for Extravehicular Activities* [36].

In our first round of testing, we compared the test subjects' performance using audio communication, the Microsoft HoloLens, and the combination of both communication methods as our three different treatment groups.

In order to analyze the data, we broke down the analysis into two major portions. In the first set of analyses, we tested for statistical significance by conducting a two-factor analysis of variance using the communication method as the treatment group and blocking the experimental design for user variation. To further determine the relative differences, we conducted a paired t-test for difference in means to see if there was statistical significance at a 95 percent confidence interval. The first set of analyses tested for statistical significance in performance time to determine if one treatment in particular was having a negative impact on performance speed.

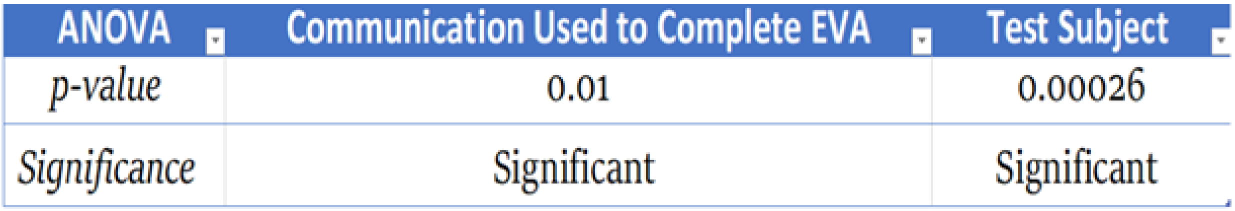
From the 30 test subjects in round one of testing, we were able to determine using an analysis of variance (ANOVA), at a p-value of approximately .01, that the communication methods were causing statistically significant differences in the average required time to perform the set of 50 tasks. Using a paired t-test for difference in means at a p-value of near 0 we were able to determine that the combination of both audio communication and the HoloLens resulted in an improved and statistically significant performance of approximately 42 seconds which is an improvement of about 12 percent. The 12 percent difference in mean performance time between the audio only communication method and the combined audio and HoloLens communication method, lead us to the conclusion that there is a slower task performance when audio is used without a spatial aid. Although smaller, there is also a statistical difference between the mean performance for HoloLens alone when compared to that of the combined HoloLens and audio method. Finally, there is no statistical difference in the mean performance between audio alone and HoloLens alone.

From our initial 30 test subjects it can be concluded that, although there is no difference between a singular auditory or visual interface, the combination of the two have a significant impact in terms of improving performance when compared to either EVA aid tested alone. The data also indicates that the combination of both interfaces has a much stronger effect in reducing performance time when compared to average user performance with audio alone. The distribution amongst the 30-test subjects suggests that there was much less variation in performance in the audio only test group indicating that the addition of a HoloLens might be creating larger variations in performance due to the steeper learning curve. With proper training and user exposure the HoloLens may improve performance even further.

A second round of analysis was also conducted on the NASA-TLX scores and the Cooper-Harper ratings in order to see if the three different forms of communication had statistically significant differences in order to determine the overall effect on user comfort. For the six TLX values and the Cooper-Harper, we once again conducted a two-factor analysis of variance and blocked for user variation checking for significance at at a 95 percent confidence level.

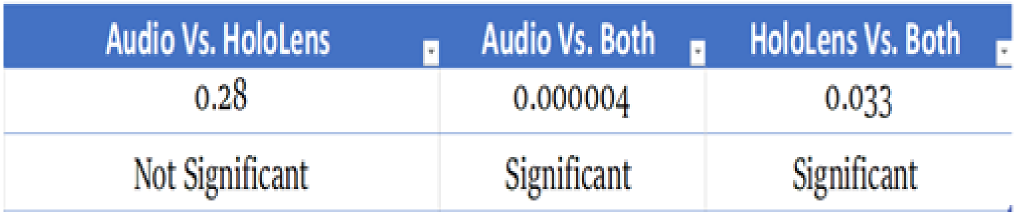
For both the NASA-TLX and the Cooper-Harper there were no statistically significant differences between the three user interfaces; however, the physical demand and mental demand portions of the NASA-TLX indicated the highest potential for statistical differences at the given significance level. If a test subject reported a “one” on the NASA-TLX, that indicated to us that the test subjects were experiencing little to no mental demand since on the NASA-TLX scale a “one” is the same as saying “very low”. One thing that we noted in our first round of testing with the mental demand is that there was an increase in lower extreme responses (1 out of 10) between the HoloLens (23 percent), the audio communication (13 percent), and the HoloLens with the audio communication (37 percent). It is also important to note that there are outliers present because everyone interprets the scale differently, and there is no calibration for this type of data. Another problem with the TLX is that subjects may have confused better performance with a higher rather than a lower number and, therefore, answered incorrectly.

Lack of differences in the remaining TLX and Cooper-Harper ratings indicates that the addition of the HoloLens as a spatial aid did not affect user comfort in a negative fashion relative to the audio communication. The combination of data from the performance and user comfort aspects of the analysis is indicating that the HoloLens is improving performance time without introducing any significant user strain.



ANOVA	Communication Used to Complete EVA	Test Subject
p-value	0.01	0.00026
Significance	Significant	Significant

Figure 9: ANOVA for round one testing.



Audio Vs. HoloLens	Audio Vs. Both	HoloLens Vs. Both
0.28	0.000004	0.033
Not Significant	Significant	Significant

Figure 10: Paired t-test for round one testing (P-values).

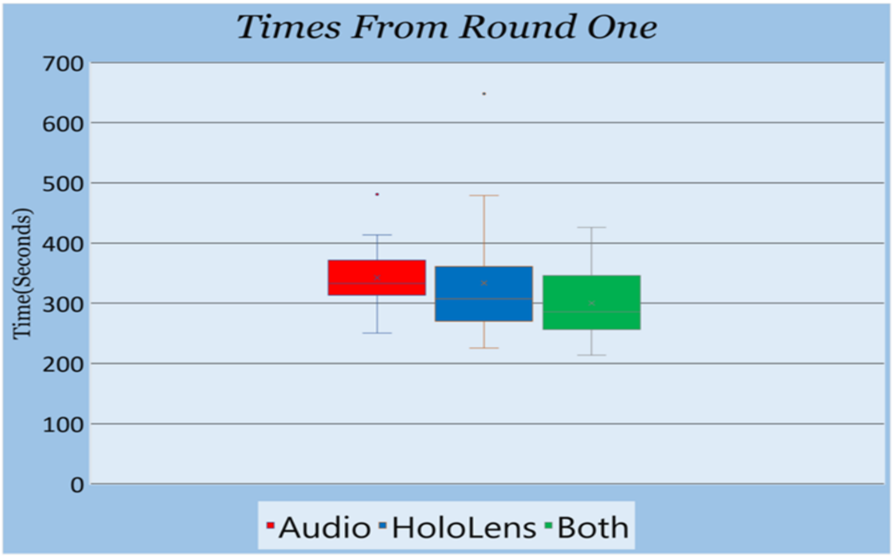
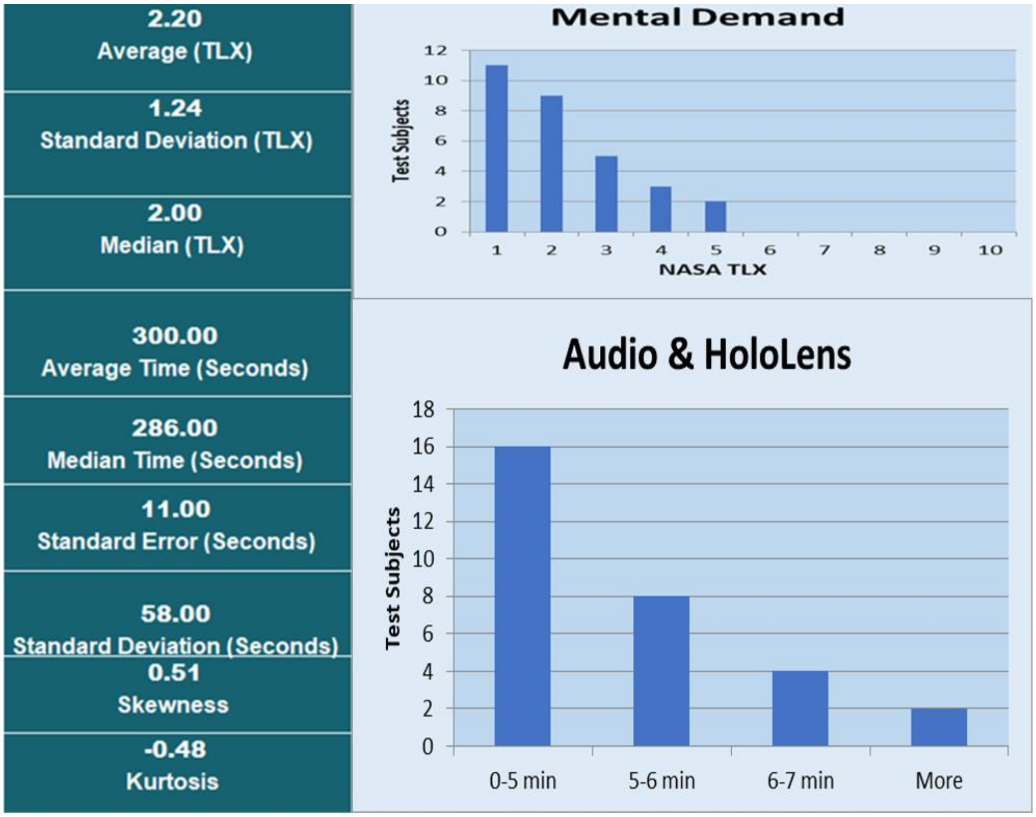


Figure 11: Round one testing time data.



2.20
Average (TLX)

1.24
Standard Deviation (TLX)

2.00
Median (TLX)

300.00
Average Time (Seconds)

286.00
Median Time (Seconds)

11.00
Standard Error (Seconds)

58.00
Standard Deviation (Seconds)

0.51
Skewness

-0.48
Kurtosis

Figure 12: Round one testing audio communication only.

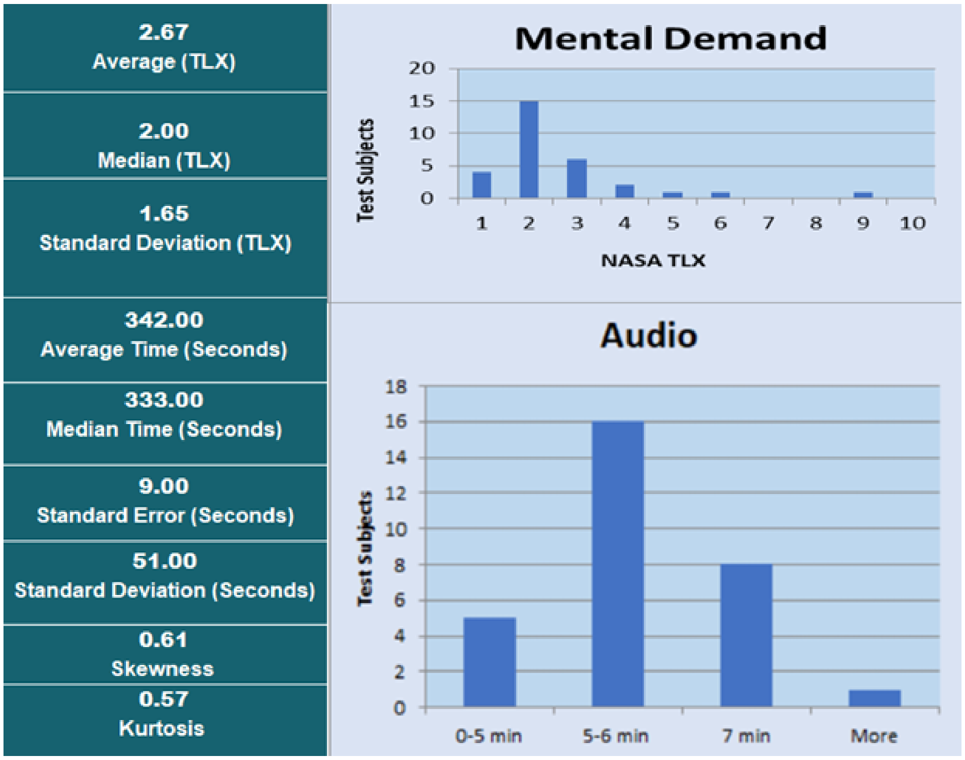


Figure 13: Round one testing audio and HoloLens.

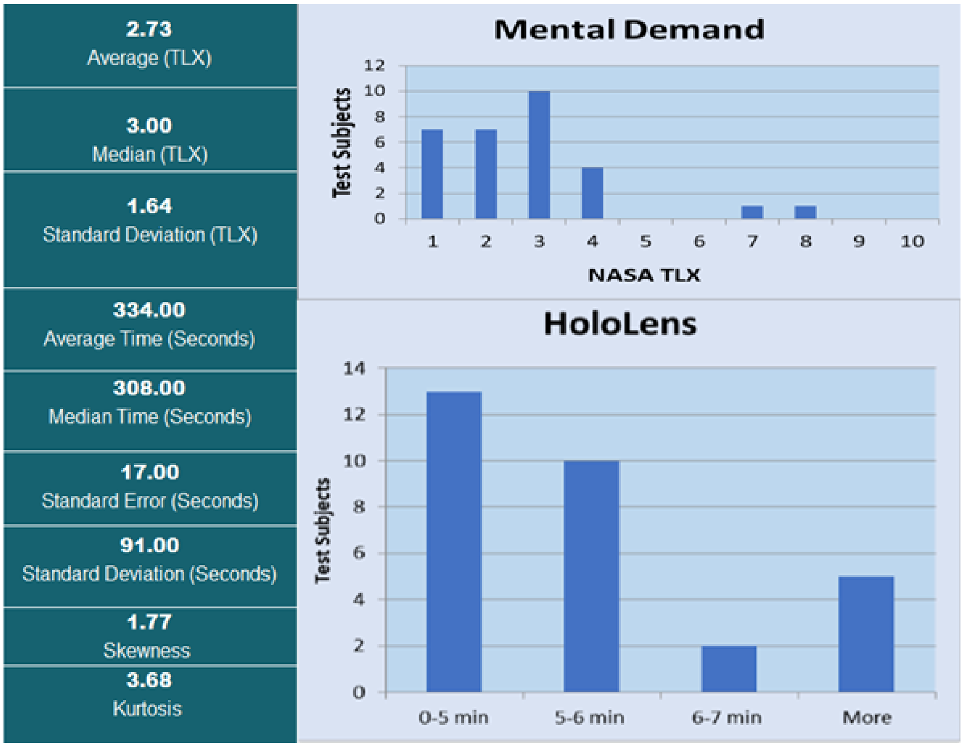


Figure 14: Round one testing HoloLens

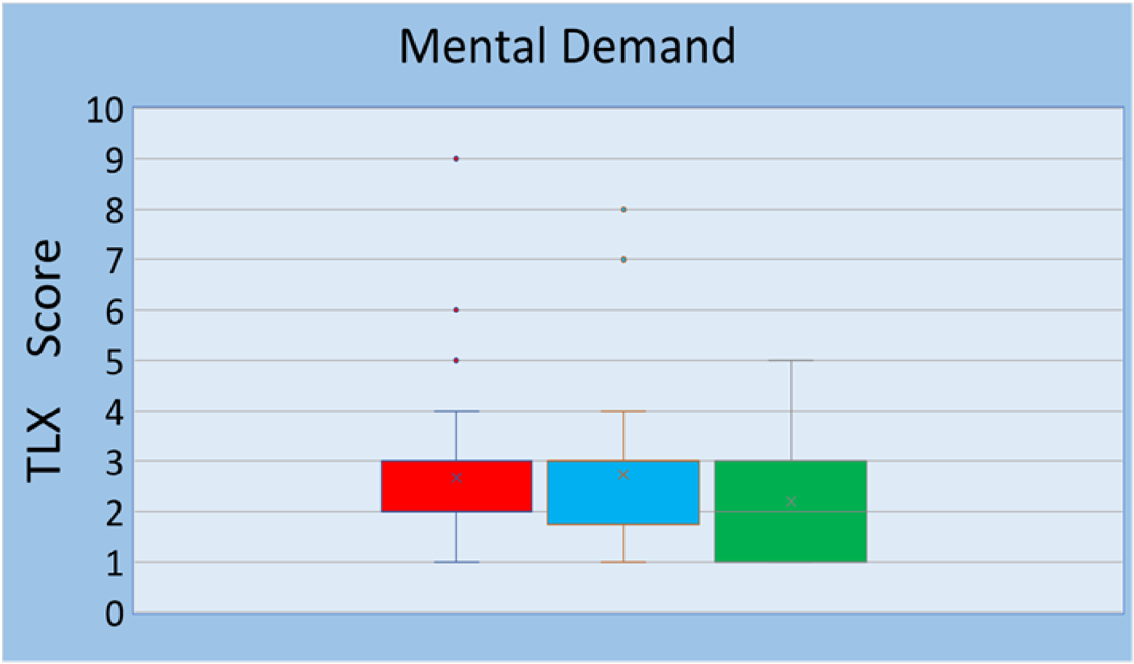


Figure 15: Round one testing NASA TLX: Mental Demand.

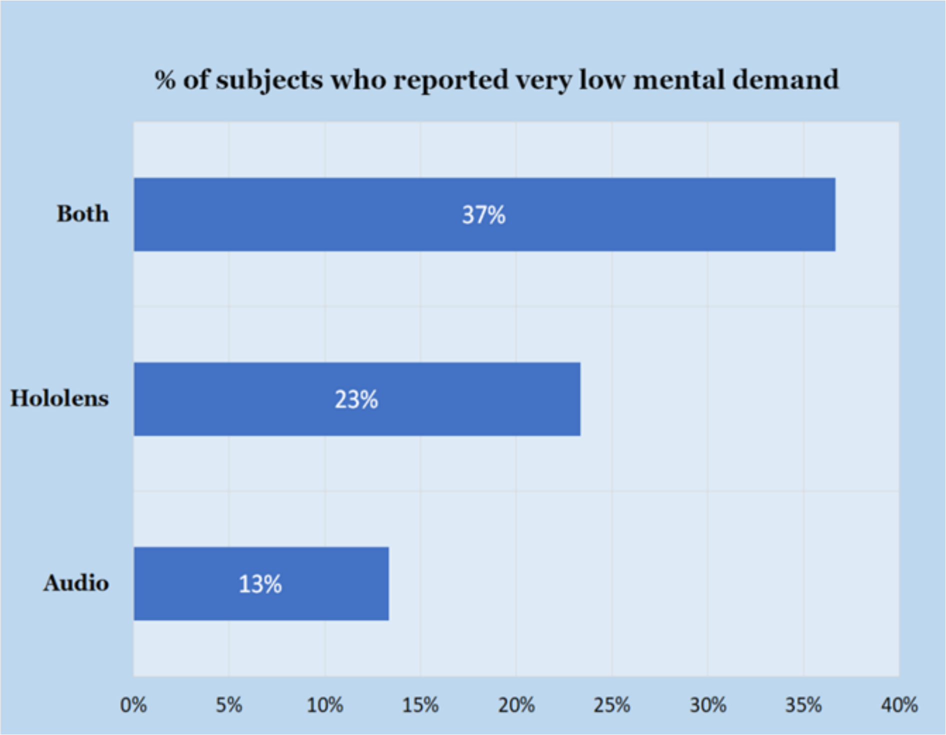


Figure 16: % of subjects who reported very low mental demand.

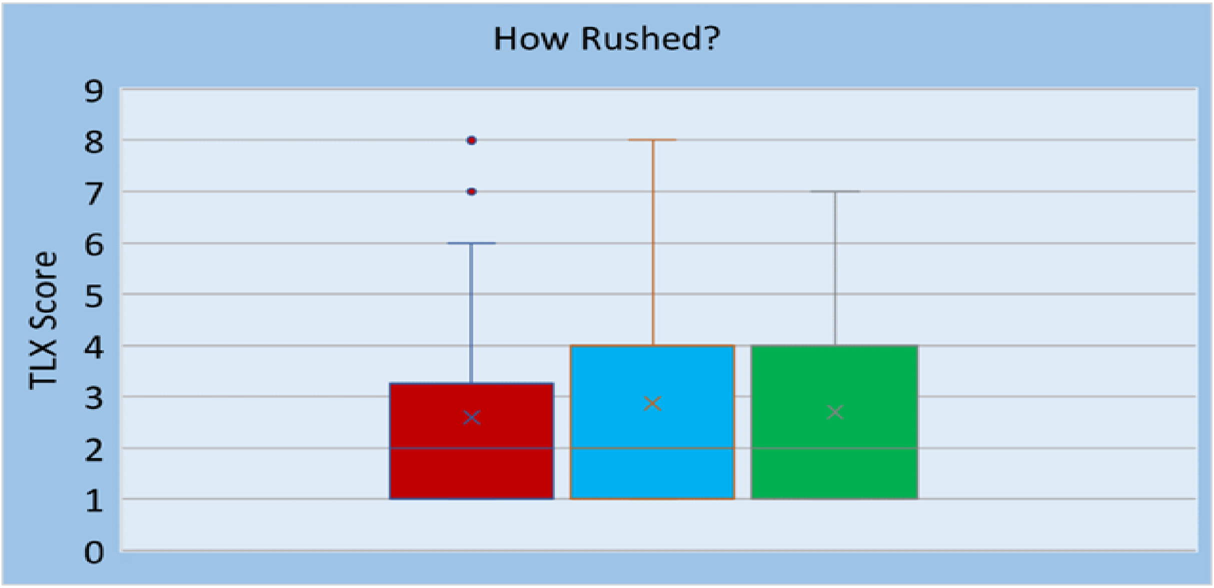


Figure 17: Round one testing NASA TLX: Temporal Demand

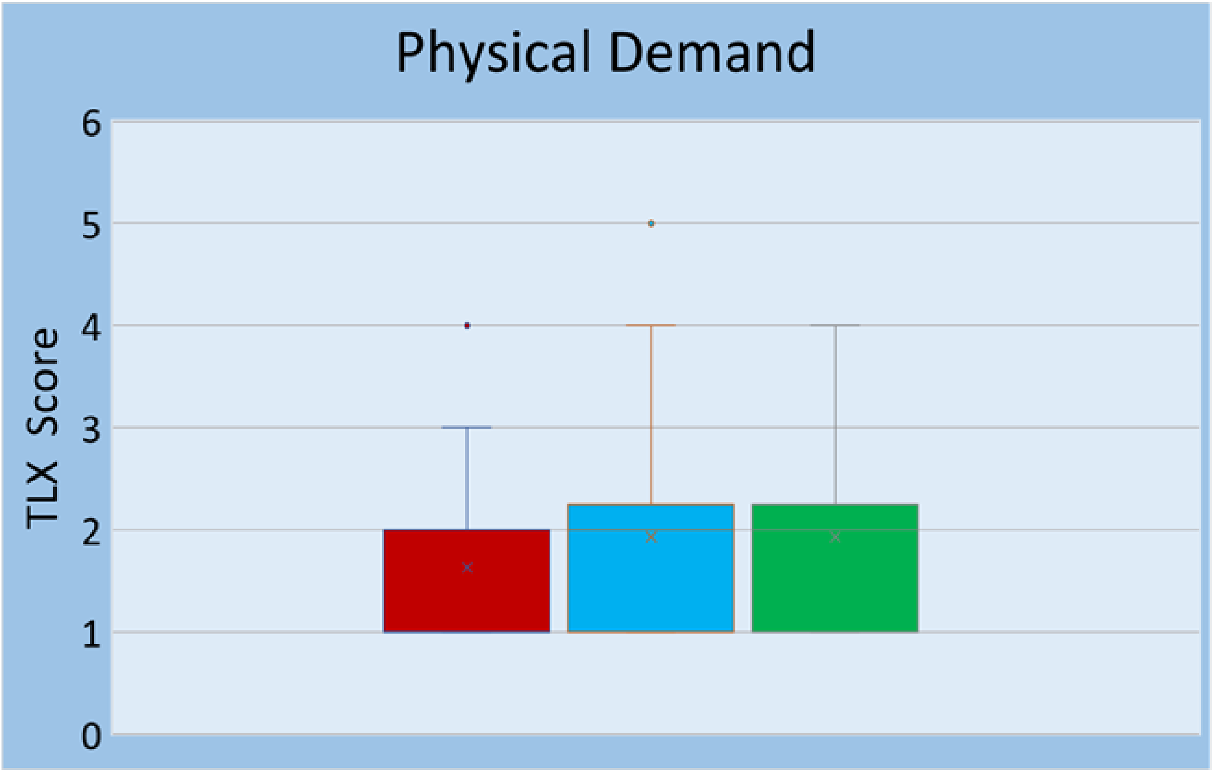


Figure 18: Round one testing NASA TLX: Physical Demand

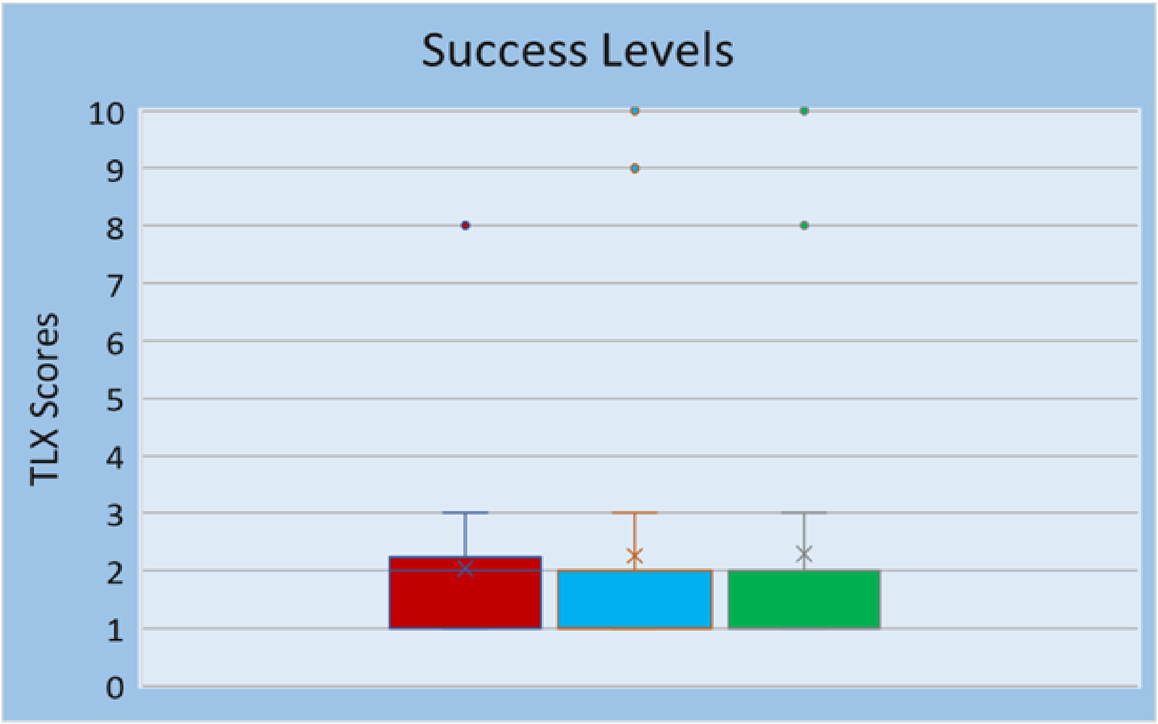


Figure 19: Round one testing NASA TLX: Performance

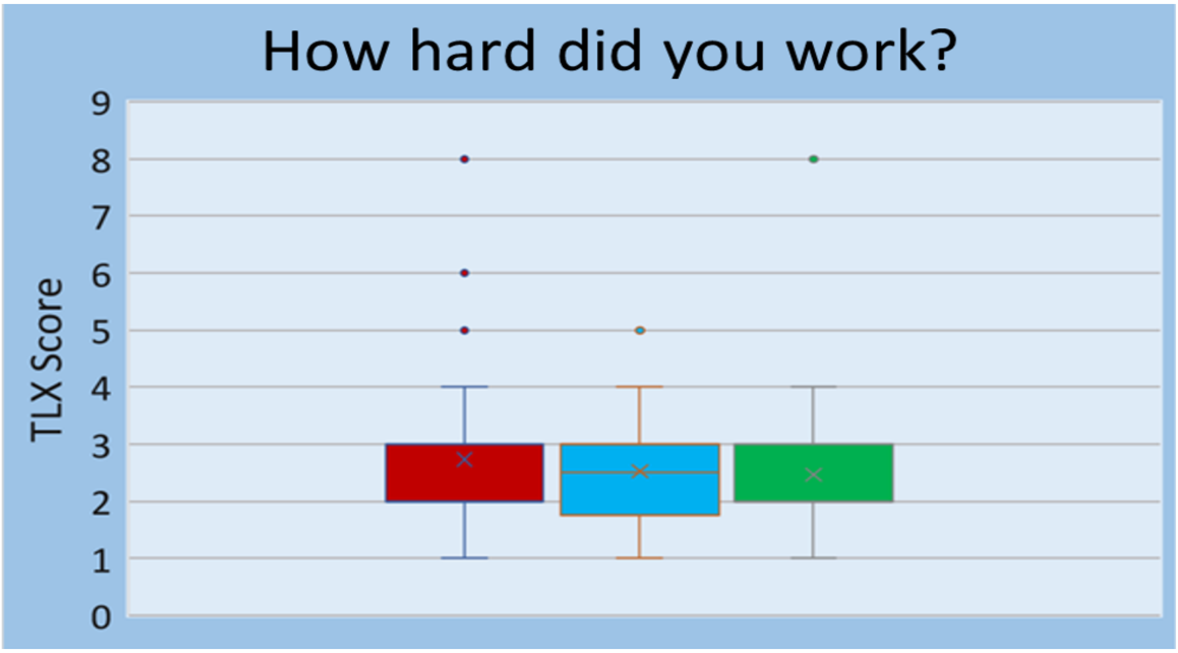


Figure 20: Round one testing NASA TLX: Effort

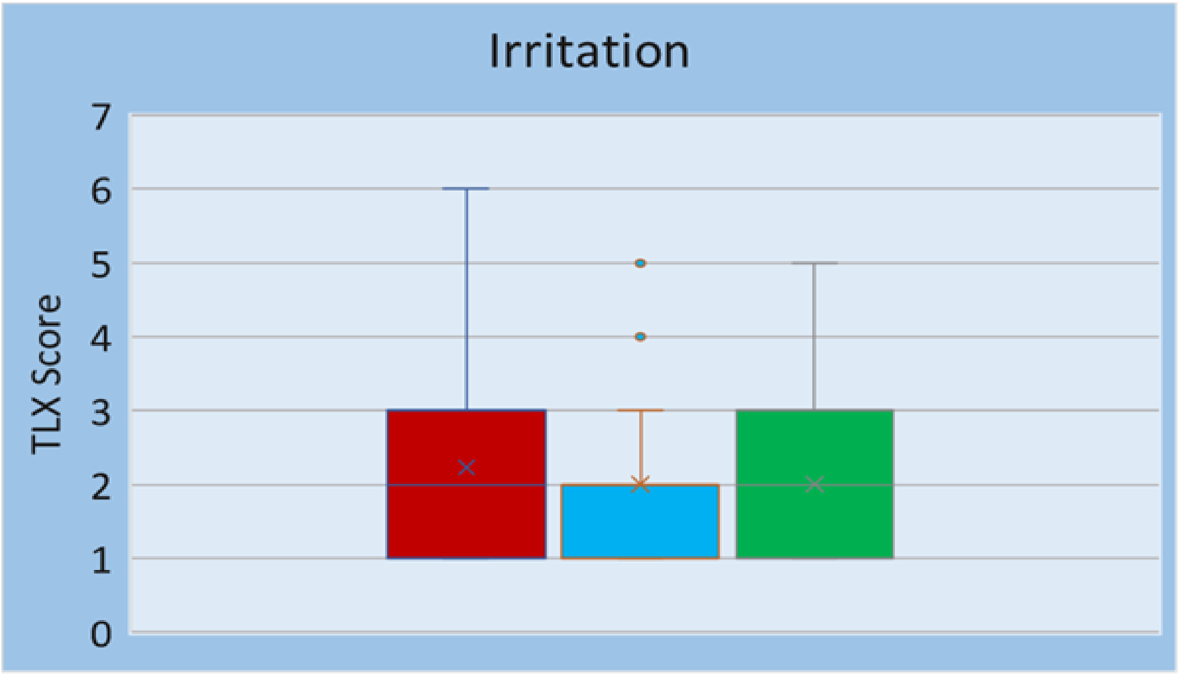


Figure 21: Round one testing NASA TLX: Frustration

The test administrators took notes on how many errors and common mistakes were made during each test. An error, for example, was the test subject being instructed to flip switch A1 but instead flipping A5. A common error was test subjects pushing button blue 1 instead of flipping switch B1 and vice versa. This happened frequently, even when the display showed a picture of the blue button when the subject was told to push it. Other less common mistakes included the wrong number on the switchboard being switched, or flipping the left and right carabiner holes. The frequency of errors was unpredictable per test and more closely varied per person. There were tests in which a test subject would make no errors and others where they made 2-3, but usually after they already completed one test, they would make fewer errors due to the learning curve factor. Another point to note with respect to the time of each test was how long the task of screwing and unscrewing the 4 holes on the grey box took. This was by far the lengthiest task and took test subjects the longest amount of time. If test subjects were unfamiliar with using a screwdriver, they would often drop screws or take longer. This problem would obviously be mitigated in a real EVA repair because astronauts would have hours of training for the task at hand and would be familiar with all of the tools that they use.

3.4 Round 2 Methodology (Prototypes/Proof of Concepts Along with HoloLens Testing)

3.4.1 Motivation

The results from the first round of testing steered the team into the next phase of testing. We wanted to collect data in a similar manner as before to test two HUD systems we designed and developed: a projector displaying information in the FOV of the wearer, henceforth referred to as our helmet-mounted display system, and an LCD screen mounted externally to the helmet, henceforth referred to as our external screen display system. Because the results from round 1 testing showed that a HUD combined with audio communication was the best for users, we divided testing into four trials: (1) audio communication only, (2) our helmet-mounted display system and audio communication, (3) our external screen display system and audio communication, and lastly (4) HoloLens and audio communication. Testing was conducted in the same manner as Round 1 using the task board with randomization of the order of tests for each test subject. The tests were again timed, and participants completed the same post-test survey with equivalent NASA-TLX, Cooper Harper, and survey and ranking questions.

With this testing methodology in mind, the helmet-mounted display and external screen display systems were designed for the ease of use while still displaying the necessary information that was designed for the round 1 testing with the HoloLens. So, they were developed to not obstruct the user's field of vision or the physical space between them and the task board, so users could still maneuver the devices on the board. Because the helmet-mounted display system was built into the design of the helmet, test subjects wore the helmet in all four tests to ensure the user's test success was not correlated to the helmet as wearing it can add a factor of stress or physical demand. A custom chest plate, which featured a ledge for the external screen display system to be mounted to, was also designed to be comfortably worn on the user's shoulders and position the helmet correctly over the test subject's head.

With all of these test design precautions in mind, testing ran smoothly for 25 subjects until it was ended due to COVID-19 restrictions at the university.

3.4.2 Hardware HUD Development

3.4.2.1 Version 1 System

The first version of the HUD design was imagined as a bespoke, entirely in-helmet setup. The goal was to produce a display directly on the helmet visor and have all of the components fit within the MX-D helmet simulator without modification to its shape.

The initial approach to assembling a HUD was buying a portable projector to scavenge components from for use inside the helmet. Originally, we intended to use two steps to this process. The first was to purchase a cheap portable projector to disassemble, modify, and bench test, and the second was to select a higher-end projector to build into the helmet. The first projector purchased was an ArtLii pico-projector from Amazon. It was fairly simple in its design and it was affordable, meaning that if it did not survive repeated disassembly and modification it could be replaced easily. It was also still small enough that it could fit inside the helmet comfortably once disassembled and rearranged, so it could be tested in the helmet if needed.

We first tested the ArtLii projector in a dark room to check the brightness and focal range of the output. Once this was done, the projector was completely disassembled. This particular unit consisted of a backlight, collimators, an LCD screen, and a lens that actually projected the image. None of these components were particularly high-quality, so the image the projector produced was not either. However, its layout was simple enough that it could be re-created inside the helmet without needing to buy another projector to remove parts from. Additionally, repackaging the components from an existing projector would require almost the same amount of effort as designing one from scratch. At this point, it made more sense to purchase individual optical components and design a completely custom projector.

Originally, a rear projection film was going to be used to boost the visibility of the projected image, since the displayed image was meant to appear on the visor. In order to test the effectiveness of this setup, two members of the team wore the helmet with the ArtLii projector held inside. They then held either a rear projection film, an anti-glare film, a combination, or neither of the two inside. They ranked the clarity, brightness, and readability of the display, as well as the readability of items outside the helmet. Ultimately, the best way to view the projected image was with no film, instead reflecting the projected image off of the transparent visor and into the viewer's eye. This setup was ranked as being the most clear by both testers, though one person ranked the image created via this method as the most difficult to focus on, while the other ranked it as the easiest. Both users found the ease of reading text outside the visor was best for this and a few other setups that did not rank as highly in other categories.

From this point, the ArtLii projector was solely used to bench test the other optical components needed for the rest of the projection system. Initially, if there were no additional lenses between the projector and the visor, the image was blurry because the visor was so close to the projection point. To test several optical lenses as additions to the optical path, one tester wore the MX-D simulator helmet while holding both the projector and a lens inside the helmet. Several testers using this method found that a double-convex lens was the best addition to the optical path, as it made the image much sharper; using a plano-convex lens also improved the image quality though not to the same degree. Using this information, the bespoke projector was developed using new collimating lenses and a slightly larger LCD screen than what was in the ArtLii unit as well as a focusing lens that would make the throw of the projector much shorter. The latter was necessary because of the short distance between the projection point and the visor.

For the bespoke projector, an LCD screen similar to the one in the ArtLii unit was purchased, but with slightly better resolution. It was also upsized from 1.9 inches diagonal in the ArtLii projector to 2.2 inches, which was the largest size that could be packaged in the helmet easily. Initially, the new LCD screen was used in conjunction with the focusing lenses from inside the ArtLii projector, but they created aberrations in the output. Different focusing lenses were purchased as a result to remove these aberrations and generate a clearer image when placed close to the LCD screen. The next issue tackled involved increasing the brightness of the LCD screen’s backlight. To increase its brightness, the LCD screen was separated from the backlight it came with and placed in front of a much brighter LED light within a custom-made reflector. The shape of the reflector was designed to mimic the reflector shape found in the ArtLii projector, but was upsized to fit the new screen and had a different interface design to allow easy swapping out of the LED backlight. The reflector was then 3D printed and lined with silver tape to increase its reflectivity. With the new backlight, the displayed image became too bright to comfortably look at, so a potentiometer was added to allow adjustments to the brightness.

Initially, the output of the bespoke projectors was going to be aimed at the acrylic visor after being focused through a collimating lens. The positioning of the projectors and the curvature of the visor were such that each projector would display an image on the opposite eye (i.e. the left projector would generate an image for the right eye, and vice-versa). Below is a diagram of the intended optical path.

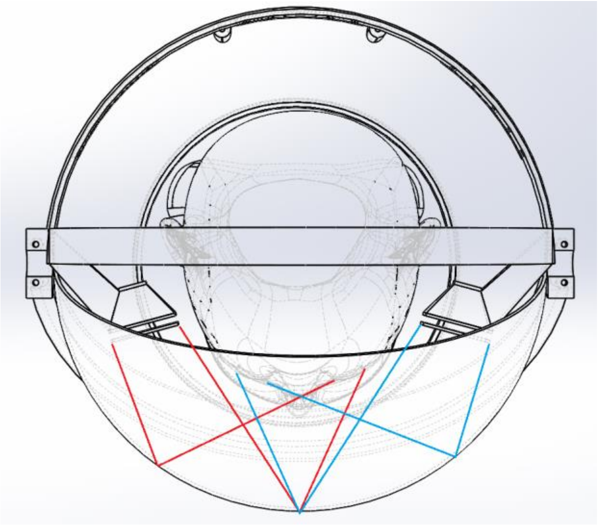


Figure 22: The proposed optical diagram for prototype version 1.

A number of problems manifested during the design process for the HUD. It was difficult for the wearer to find the image on the visor - they could only see it if they held their head in one specific position. The image was also far too unfocused to see clearly or read any text from, due in part to the curvature of the visor. This also resulted from the issue that aligning both projectors so they would actually project clearly to both eyes at the same time was exceedingly difficult. The placement of the projectors in the helmet also meant that they came very close to the wearer's head, which would limit their head motion. Overcoming the optical issues alone would have required much more lensing, and packaging all of the new components would require a redesign of the helmet shape.

Since so much work was required to overcome the first prototype's problems, we decided to scrap this design and move to something simpler. The design would now use more off-the-shelf optical components as well as reshape the helmet to allow more room. The next phase of prototype development would stem from this decision.

3.4.2.2 External Screen Display

3.4.2.2.1 Why use a screen?

Our experimentation with an LCD screen during the first phase of development was useful for one of our new approaches: an external screen mounted outside the helmet. An advantage of this approach was that there were small screens available that would still be able to adequately display instructions, which was helpful for obscuring as little of the FOV as possible. In addition, a screen could be easily mounted outside the helmet but still within the FOV of the user, portraying information while adding a minimal amount of burden to the user. This also allowed us to compare the external screen display to the helmet-mounted display and HoloLens. Another benefit was that the screen could be maneuvered, unlike some other displays. This allows the user to position the screen, and therefore the instructions, where it was most comfortable while still allowing effective use of the screen. All these traits made an external screen system a desirable setup to test with our software.

3.4.2.2.2 Mounting

Due to limited space within the helmet, the screen needed to be mounted outside the helmet to increase comfort for the user. A commercial swivel holder with a mounting clamp, commonly used for cell phones, was purchased and an adapter was made to connect the holder to a ball and socket joint. Then, a 3D printed casing (shown in **Figure 23**) was made to hold the screen and connect to the adapter. The mounting clamp from the swivel holder clamped down onto the chest piece of the shoulder mount, which will be described in the secondary components section. The entire system could be moved to any location in front of the visor that is comfortable for the user to see while remaining out of their main FOV.

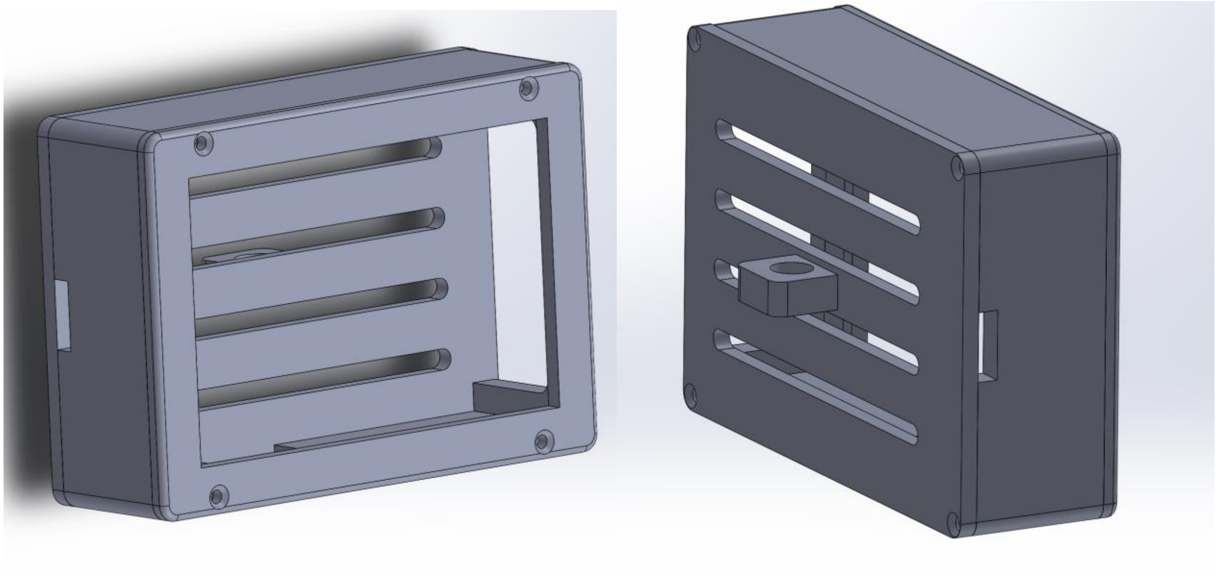


Figure 23: 3D printed screen case.

3.4.2.3 Helmet-Mounted Display

3.4.2.3.1 Why Use a Projector?

The Microsoft HoloLens used in the preliminary testing is an AR HMD, as it creates an image suspended in space in front of the user's direct line of sight while being worn directly on the head. One might question the motivation to produce a helmet-mounted display to test alongside the HoloLens; however, it is important to recognize the unfavorable nature of using a HMD inside of a spacesuit helmet. If problems were to arise with the display during an EVA, the astronauts need a way of manually deactivating or removing it without accessing the internals of their suits. This is not possible if the display and its controls are mounted on the user's head within the pressurized helmet. Therefore, it is important to develop a helmet-mounted display that is integrated into an astronaut's helmet rather than mounted directly on their head. We achieved this by using a projector mounted above the back of the astronaut's head in tandem with a diffuser and teleprompter glass placed in front of the helmet's visor. The placement of these components avoids any possible interference with the astronaut's limbs and does not directly impede their vision, all while providing external access to both the projector and teleprompter glass in case of an emergency. The details of the optical components and design process for an unpressurized, 3D printed prototype will be discussed in the coming sections.

3.4.2.3.2 Projector Selection

The process for selecting the projector was a design trade study. Six different off-the-shelf projectors were compared based on different metrics; using these metrics, we then selected two projectors that were selected and purchased. The metrics that the team used to compare the projectors were cost, size, weight, what ports it had (i.e. High-Definition Multimedia Interface [HDMI] or Universal Serial Bus [USB]), battery life, and throw distance (if specified), in order of priority. In the prototype, we decided to use the Vamvo Ultra Mini Portable Projector because it met our price requirements of being below $250 and had the right dimensions for our compact design. It worked flawlessly during prototyping and testing.

3.4.2.3.3 Optical Components

Several optical components were needed in order to make the projection system function correctly. Similar to a teleprompter or the HUD in a modern car, the in-helmet projection system uses a diffuser and reflective glass to generate a clear image and reflect it back to the user’s eye.

The first component after the projector is a diffuser, which acts as a projection screen of sorts. The wearer sees a reflection of the image displayed on the diffuser when it is in place rather than just seeing the source from the projector [37]. The diffuser used was one of the rear projection screens purchased for phase one of the HUD development, since they were already on-hand and produced a very clear image when paired with the new off-the-shelf projector. For mounting, the diffuser was placed inside the helmet on the back edge of the visor, so that it sat just behind and above the wearer's head. This placement coupled with the location of the projector in the helmet and the projector's throw ratio generated an image that was approximately four inches across.

The next component in the optical path is a piece of teleprompter glass. This was selected because it is reflective enough to allow the wearer to clearly see the image on the diffuser, while still being transparent enough to see through if no image is displayed [37]. The ratio of transmittance to reflectance was varied slightly during development. Originally, glass with 40 percent of the incident light reflected back to the wearer was used, but this was later reduced to 30 percent because the 40 percent reflectance caused the wearer's face and the inside of the helmet to be faintly visible on the glass when in use.

Originally, we planned to use a mirror to redirect the diffuser's image before it struck the teleprompter glass. This layout would place the glass directly in front of the wearer's face and the mirror above their head to turn the image downward toward it. However, it was decided that the glass was better placed above the wearer's head where it would not obstruct their field of view at all, making the mirror no longer necessary.

3.4.2.3.4 Design Iterations

Our helmet-mounted display system required an iterative approach involving repeated bench testing. It was important to ensure that the placement and orientation of the projector, diffuser, and teleprompter glass rendered a readable and properly sized image for the user, and that the image was placed in a convenient location that did not interfere with the user's ability to complete the tasks.

To accomplish this, a test structure was designed and built using 80/20 aluminum extrusions. The structure allowed for the projector to be mounted such that its vertical and horizontal position relative to the helmet could be adjusted. Additionally, the angle between the projector and the horizontal could be adjusted. The user testing the placement of the components was able to insert their head into the bottom of the helmet as the structure cantilevered the assembly over a table or other flat surface, as shown in **Figure 25**. The diffuser was mounted rigidly to the top edge of the visor as seen in the image of the test setup. The bottom of the diffuser sat above the top of the user's head. The teleprompter glass was held by the user and its angle, as well as vertical and horizontal positions relative to the helmet, were adjusted. These general placements of the optical components were determined to be the most likely to provide a clear, undistorted image, while also being least likely to interfere with the user.

Figure 24: Projector test setup.

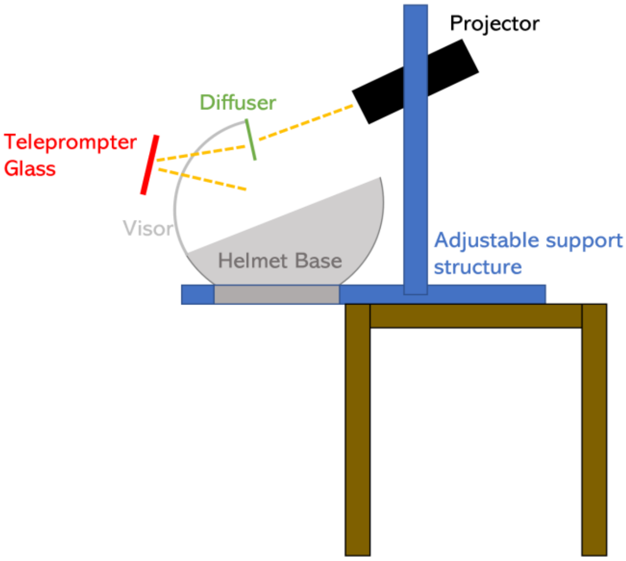


Figure 25: Projector test setup diagram.

The bench testing consisted of various members of the team placing their heads inside the helmet and adjusting the placement of the projector and teleprompter glass until a clear image was displayed on the glass, and the glass was not in a position that would distract the user from their immediate surroundings. It was important that the projector was not placed too far away from the center of the helmet, so as to mitigate any difficulties with balancing the weight above the user's head in what would be the eventual final assembly. The manual zoom on the projector was used to focus the image regardless of the physical placement of the projector. This made for very easy and repeatable adjustments. When a final placement was agreed upon, the angle, horizontal and vertical positions of the projector and teleprompter glass were measured relative to identifiable points on the helmet. The measurements are shown in **Figure 26**.

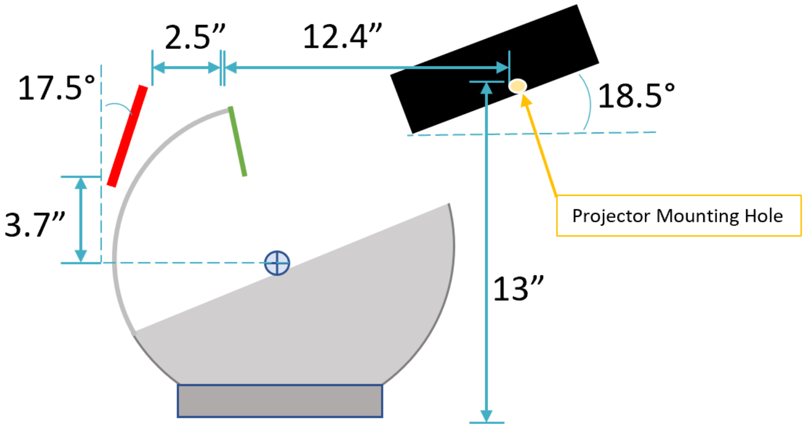


Figure 26: Projector and teleprompter glass placement.

The next step was to design a housing for the projector and teleprompter glass that would allow for easy access to the projector ports and controls while also ensuring the proper placement of these optical components. This required modifying the top half of the 3D printed helmet to have a cutout in the back as shown in **Figure 27**.

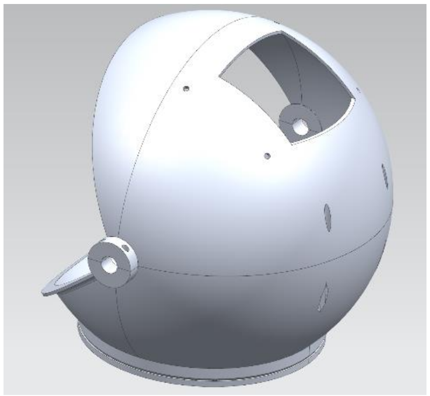


Figure 27: Top of helmet cutout.

The projector housing consists of two parts. The bottom part is shown below on the left. It interfaces with the helmet via bolts and nuts through two mounting holes, and allows for the projector to be secured in place using a set screw through the flat portion of the housing which mates to the mounting hole in the bottom of the projector. The bottom portion of the housing also features a large cutout to allow for access to the projector's manual focus, HDMI and USB ports, and power port. The top portion of the housing is fastened to the helmet similarly and rests flat on the bottom portion.

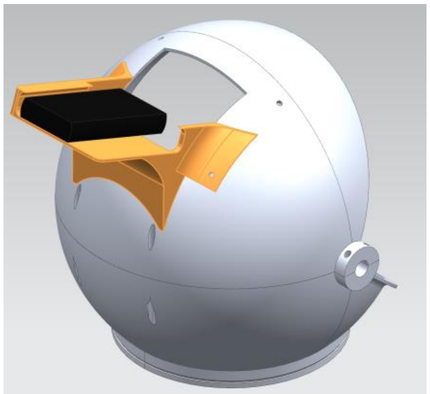


Figure 28: Projector housing (bottom).

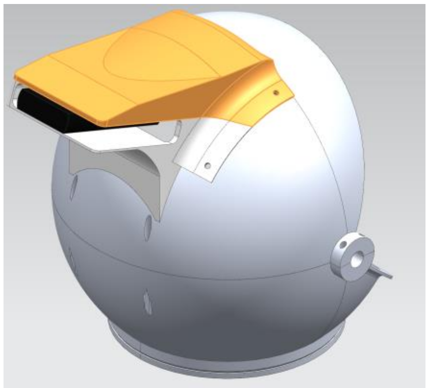


Figure 29: Top half of projector housing.

The mount for the teleprompter glass was secured to the top of the helmet using Velcro. The Velcro allowed for the placement of the glass to be adjusted and fine-tuned as needed. It featured slots in which the glass could be inserted from the top to be secured in place at the specified angle and height determined. The teleprompter glass mount is highlighted in **Figure 30** and **Figure 31** below.

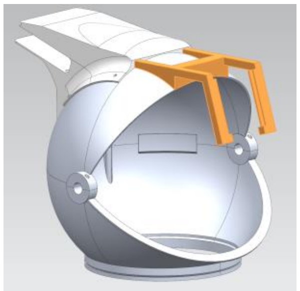


Figure 30: Teleprompter glass mount.

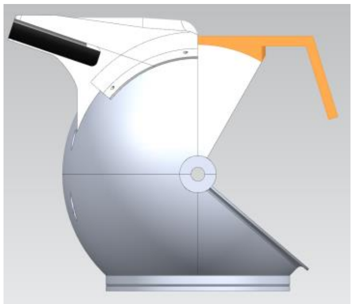


Figure 31: Teleprompter glass mount (Side).

The final computer-aided design assembly of the projector HUD helmet system is shown in **Figure 32**. The 3D printed prototype used for testing is shown in **Figure 33**. This image also includes the visor as well as teleprompter glass, and the helmet is shown resting on the shoulder mount. The shoulder mount, along with some other secondary components, will be discussed further in the next section.

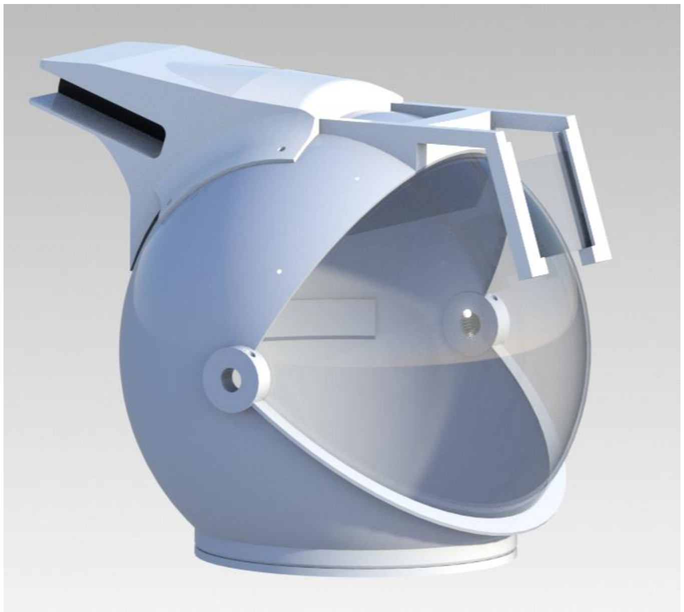


Figure 32: Render of full helmet-mounted display prototype.

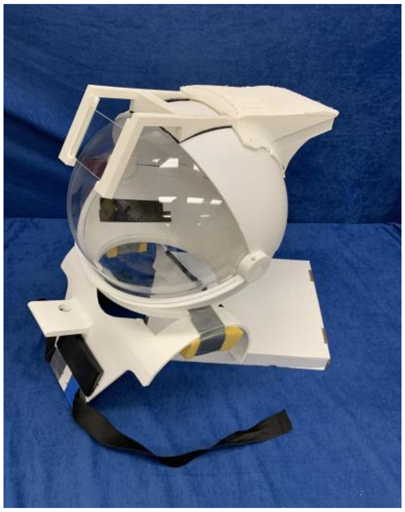


Figure 33: Helmet-mounted display prototype with shoulder mount.

3.4.2.4 Secondary Components

3.4.2.4.1 Shoulder Mount Development

In order to test the different systems with the helmet, a shoulder mount needed to be developed to support the helmet on the user. It was important that the shoulder mount had a compatible interface with the helmet, was comfortable for different users, and did not restrict the user's motion. To ensure a compatible interface, the team used an existing helmet stand that consisted of a simple ring in which the helmet could sit in as a baseline. We went through two iterations in the development of this shoulder mount.

The core of Iteration 1 is shown below in **Figure 34**. This iteration used the initial simple ring and added material to the shoulders and chest. Additionally, foam padding and Velcro straps were added to increase comfort and stability. We found that even with the foam padding, the mount did not support the helmet enough, and that the addition of Velcro straps around the users' arms restricted their motion.

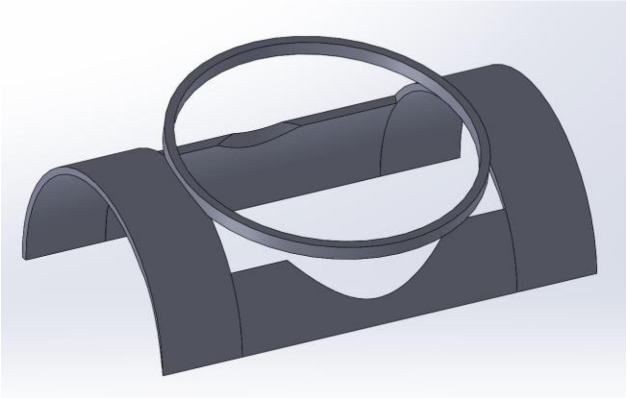


Figure 34: Shoulder mount Iteration 1.

Iteration 2 of the mount, shown in **Figure 35,** was developed to solve some of the issues noticed with Iteration 1 and to account for additional changes to designs of other components. Material was added to create a chest piece, solving problems encountered with instability in Iteration 1. An extrusion from the chest piece was included to allow the external screen display to easily clip onto the mount, and a hole was included on that extrusion for wires to be organized. Additionally, edges were rounded to avoid injury and stress concentration at sharp corners. Not shown in the figure, foam padding was added in select areas for comfort and an optional Velcro strap around the chest was included if users felt they needed more stability.

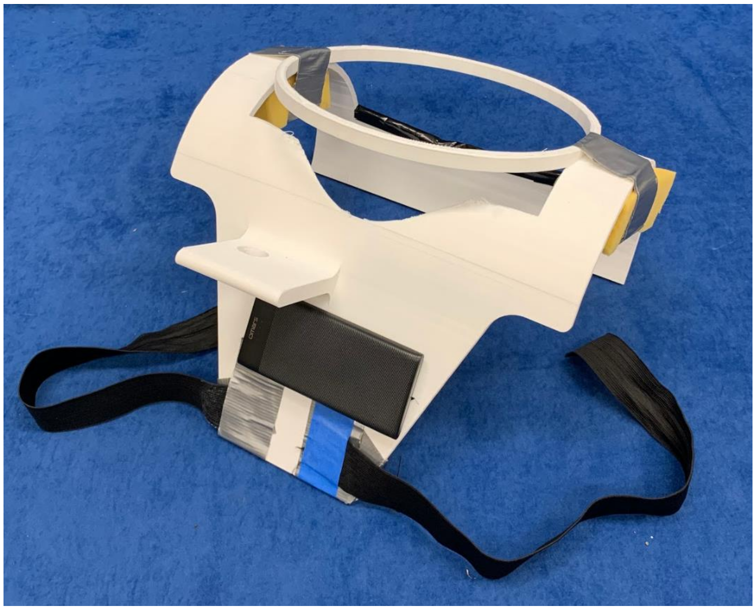


Figure 35: Shoulder mount Iteration 2.

3.4.2.4.2 Supporting Electronics

A few supporting electronic components were used for the second phase of prototype development. These were mostly to provide communication between the display systems and the test computer running the software, as well as to provide power to the displays. Originally, we intended to host the software used by the displays on a Raspberry Pi. This would allow the helmet to carry its own on-board software and not require an external computer for control. However, this approach was abandoned because the external computer was still required for HoloLens testing, and it made more sense to simply continue using it for the other tests as well. Instead, the displays on the helmet were connected to the test computer via a wireless HDMI emitter plugged into the HDMI port on the computer and a wireless HDMI receiver plugged into the helmet-mounted display or external screen display. Each system would then act as a second screen for the computer that would show whatever the computer was displaying. To power the HDMI emitter/receiver, a pair of rechargeable lithium battery packs were purchased. These batteries also powered the screen in the external screen display, but not the projector in the helmet-mounted display system since that unit had its own power supply.

The initial test plan called for the wearer to use headphones to hear the audio instructions from the testers, but during preliminary testing these were deemed unnecessary and were dropped from the main test procedure.

3.4.3 Round 2 Testing Process

For the second round of testing, the process and procedures followed were largely the same as those for the first with a few exceptions. Because of their success, simplicity, and for the purpose of consistency, the simulated EVA task board, and therefore the 50 mechanical tasks participants were asked to complete, was exactly the same. In addition, the same recruitment process was used to gather adults of the same population pool and with the same requirements. Each test again had a randomized order of tasks, and the order of the trials was also randomized to prevent the effect of the learning curve. The pre-participation and post-tests surveys used to gather data were essentially exactly the same as those from the first round, with only very minor superficial alterations to accommodate the new tests and make slight corrections.

As mentioned, though, there were a few differences for the tests in the second round. The main, and naturally most important, one was the fact that the forms of communication for each trial were new this time around. Given our new technologies and motivations, the testing now consisted of four trials for each participant: one with audio communication alone, one with audio and the Microsoft HoloLens running our software HUD, one with audio and our helmet-mounted display prototype, and one with audio and our external screen display prototype. Based on the results of the first round of testing, it was extremely clear that every metric of performance and satisfaction was higher for trials combining audio and visual stimuli as compared to the visual forms alone, which is why only these four tests were conducted, ignoring trials with the HoloLens or either of our prototypes by themselves. To go along with this, another significant change was that subjects wore our team’s simulated spacesuit helmet for all four tests so that there was consistency between each trial and so any effect the helmet itself (and not the actual communication methods) may have had in efficiency, ease of use, or comfort would be mediated.

The final key difference was that the sample size of the data for the second round of testing had 25 subjects, as opposed to 30 for the first. Initially, the goal was for the team to test on and obtain data for at least 30 participants for the same reason as in the first round, and potentially look at testing as many as 50 if possible to further attain statistical significance with an even larger sample size. However, based on time limitations and the introduction of major lifestyle changes and the cancellation of all on-campus education and research activities due to the COVID-19 pandemic, our team ended up only being able to test slightly less than 30 subjects. After eliminating some errant data, the final data set for the second round of testing had a sample size of 25. Although this is less than that of the first round and is less than the team would have desired, it is still enough for statistically significant findings to emerge and is an achievement given the circumstances.

3.5 Round 2 Results and Discussion

3.5.1 Results

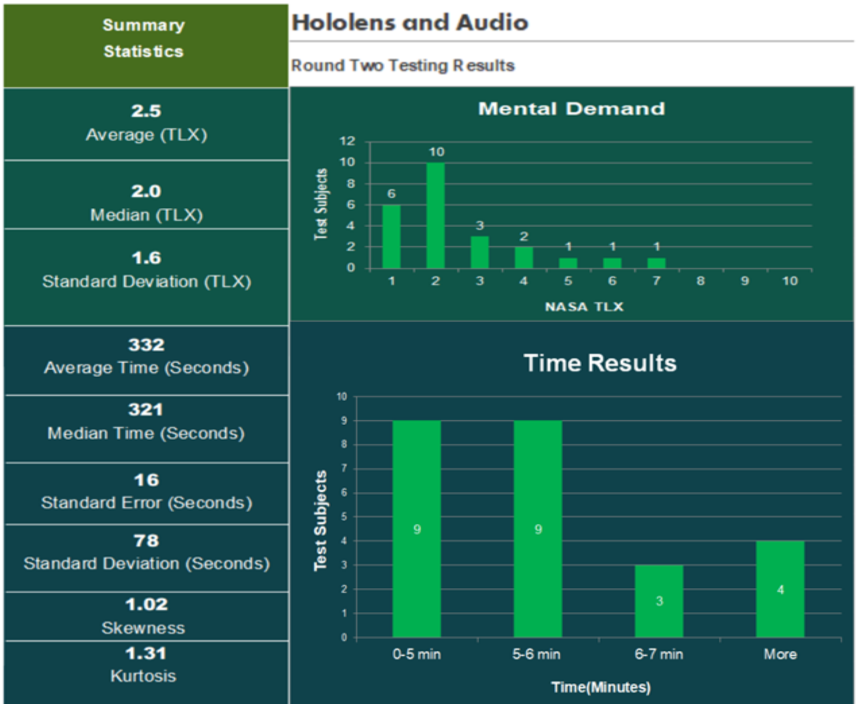
3.5.1.1 Efficiency (Times)

Saved time over the course of an entire EVA can improve the safety of an astronaut by reducing the hours spent on EVA. In our second round of testing, we assessed time efficiency and other parameters using four treatment groups: audio communication only, the Microsoft HoloLens with audio, a helmet-mounted display with audio, and an external screen display with audio. Similar to the first round, we collected the times of each performance and completed statistical analysis on the means of each treatment group. This could help us understand if the combination of audio and visual communication would help test subjects perform the EVA tasks using our new sample of 25 test subjects. Since the first round of testing, we had prototyped a helmet-mounted display and an external screen display that could offer the same benefits as the HoloLens, but be mounted externally to avoid a system that has to be mounted directly onto the user. Round two of testing was different from round one since the data could demonstrate if the location of the visual display had an effect on the time.

Similar to the first round of testing, we used ANOVA, as well as a paired t-test to understand how the different treatment groups affected the average performance across the 25 test subjects. For round two testing, it is important to understand the summary statistics in a few main areas before we compare the sample means. In order to understand the data, we will talk about the measures of central tendency, the measures of the spread of the data, and how normal the sample data is compared to the overall population. Similar to the first round, we selected a sample size close to 30 because of the central limit theorem. As written in a publication in the National Center for Biotechnology Information,

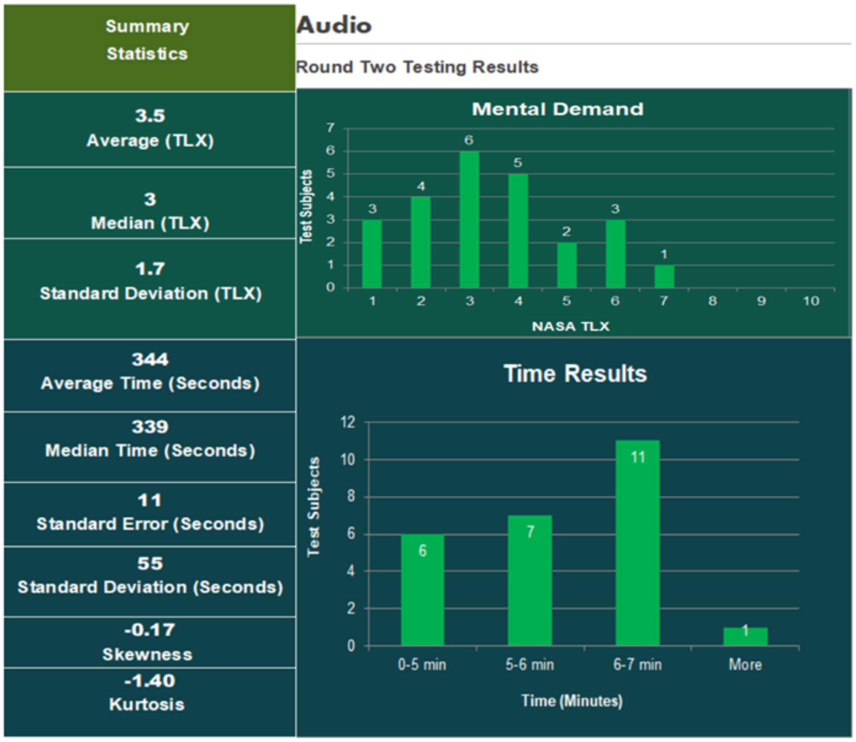
> "If the sample size is 30, the studentized sampling distribution approximates the standard normal distribution and assumptions about the population distribution are meaningless since the sampling distribution is considered normal, according to the central limit theorem. Therefore, even if the mean of a sample of size > 30 is studentized using the variance, a normal distribution can be used for the probability distribution" [38].

Each test subject was also assigned a randomized order for the type of communication they would use first, second, third, and fourth in an effort to prevent learning from affecting the time results. In round two of testing, 6 people used audio communication only first, 7 people used the HoloLens first, 6 people used the helmet-mounted display first, and 6 people were given the external screen display first. Based on those numbers, we can be more confident that average times were not significantly affected by the order, since the odds of getting similar orders using a random number generator over 25 test subjects is fairly low and the amount of time that each communication was given first in the sequence was uniform (other than the HoloLens which had 7 test subjects rather then 6). Even in the first round of testing, 9 people received audio communication only first while 7 people received the HoloLens and audio first. The distribution of times and summary statistics for round two testing can be seen in figures below.



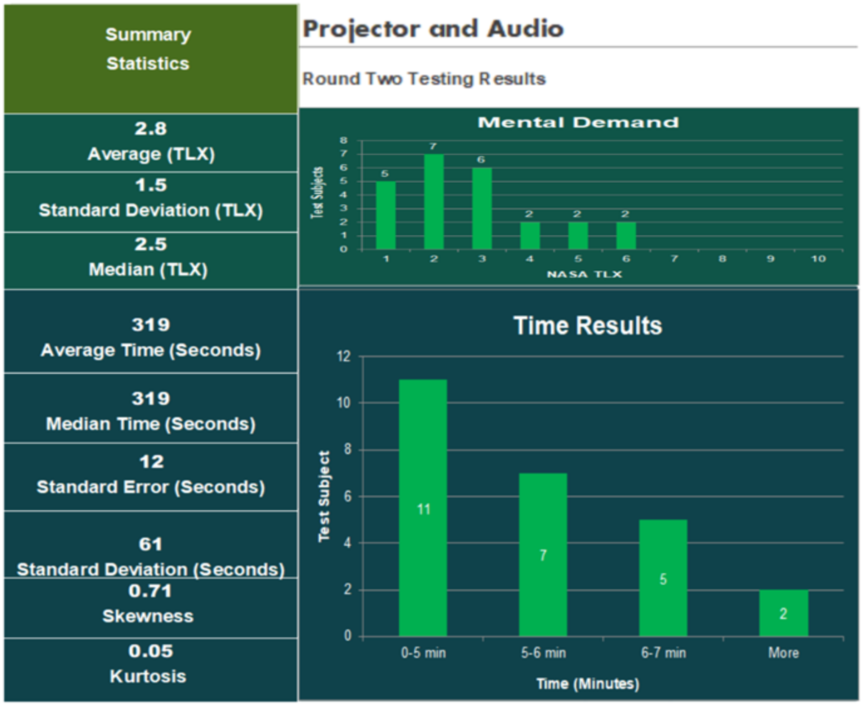
Summary Statistics	
2.5	Average (TLX)
2.0	Median (TLX)
1.6	Standard Deviation (TLX)
332	Average Time (Seconds)
321	Median Time (Seconds)
16	Standard Error (Seconds)
78	Standard Deviation (Seconds)
1.02	Skewness
1.31	Kurtosis

Figure 36: Round two testing: HoloLens



Summary Statistics	
3.5	Average (TLX)
3	Median (TLX)
1.7	Standard Deviation (TLX)
344	Average Time (Seconds)
339	Median Time (Seconds)
11	Standard Error (Seconds)
55	Standard Deviation (Seconds)
-0.17	Skewness
-1.40	Kurtosis

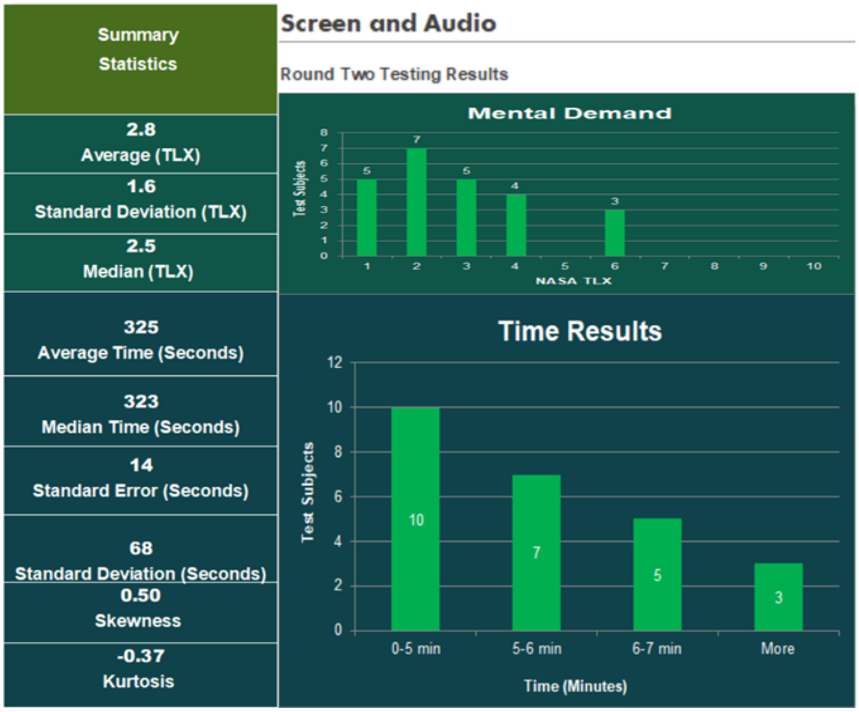
Figure 37: Round two testing: Audio



Summary Statistics

2.8	Average (TLX)
1.5	Standard Deviation (TLX)
2.5	Median (TLX)
319	Average Time (Seconds)
319	Median Time (Seconds)
12	Standard Error (Seconds)
61	Standard Deviation (Seconds)
0.71	Skewness
0.05	Kurtosis

Figure 38: Round two testing: Projector



Summary Statistics

2.8	Average (TLX)
1.6	Standard Deviation (TLX)
2.5	Median (TLX)
325	Average Time (Seconds)
323	Median Time (Seconds)
14	Standard Error (Seconds)
68	Standard Deviation (Seconds)
0.50	Skewness
-0.37	Kurtosis

Figure 39: Round two testing: External Screen

We were able to determine that our kurtosis values for time were within a reasonable range since “a kurtosis value of +/-1 is considered very good for most psychometric uses, but +/-2 is also usually acceptable” [39]. The highest skewness we saw in the time results was around 1.02 (HoloLens), and based on research “values for acceptability for psychometric purposes (+/-1 to +/-2) are the same as with kurtosis” [39].

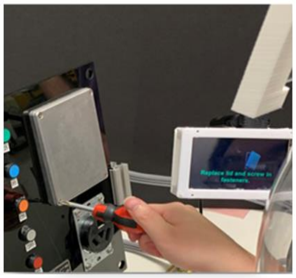
Starting with the measures of central tendency for the audio, projector, and screen we can see that average time is reasonably close to the median time, but for the HoloLens the 11 second difference shows that it was a little more affected by outliers compared to the other three, but not significantly. Comparing the averages, we see that compared to the audio, the test subjects performed better with the use of any form of visual or augmented display. The improvement in time in seconds compared to the audio as well as the improvement as a percentage of the average time taken to complete the task using audio communication can be seen in **Figure 40**.

	Audio	***Hololens***	***Screen***	***Projector***
Mean (sec)	344 sec	332 sec	325 sec	319 sec
Mean (min)	5.7 min	5.5 min	5.4 min	5.3 min
Improvement over Audio (sec)	--	12 sec	19 sec	25 sec
% Improvement over Audio	--	**3%**	**6%**	**7%**

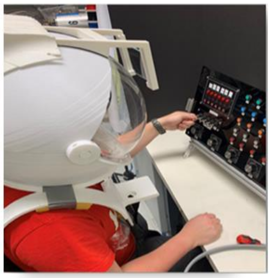
Figure 40: Round 2 testing average times

It was interesting that with the addition of the spacesuit on top of the HoloLens, we are only seeing an average difference of 12 seconds, which is smaller compared to the 42 second difference we saw in round one testing. Comparing the HoloLens, screen, and projector to the audio using a paired t-test on the sample means, we are able to find that the time improvements seen with the projector (P-value one tail .045) and screen (P-value one tail .088) were statistically significant within a 90 percent confidence interval. We are seeing decreases in performance that are around six to seven percent, but we expect that the percentages would be even greater, since the difficulty and length of EVA would be much greater than the 50 tasks we had for our test subjects.

t-Test	Audio	Hololens
Mean time (sec)	344 s	332 s
P(T<=t) one tail	**0.217**	



t-Test	Audio	Screen
Mean time (sec)	344 s	325 s
P(T<=t) one tail	**0.088**	



t-Test	Audio	Projector
Mean time (sec)	344 s	319 s
P(T<=t) one tail	**0.045**	

Figure 41: Round 2 testing t-Test results.

Conducting a two-factor analysis of variance on just the visual communications blocking for the user and the type of communication, we found that at a p-value of .64 there is not a statistically significant difference between the HoloLens, our external screen display, and our helmet-mounted display. This helped us to see that the location of the visual display did not have a major effect on the average times.

Comparing the standard deviations of the time also shows that the screen and the HoloLens had higher standard deviations than the audio and projector, which shows that those displays may have allowed some test subjects to perform much better than others possibly because of experience using something that was similar. The standard errors we were seeing ranged from 11 to 17 seconds, which are within a reasonable range given that our sample size was on the smaller side of n = 25.

3.5.1.2 NASA-TLX Score Analysis

The NASA-TLX asks users six different questions, which they can score on a scale of 1-10 depending on the demand that the test subjects experienced. In order to assess the effect of the communication treatment groups, we looked at an average score across the six questions. Across the six questions the averages showed us that the mental demand experienced by the test subject across the four different types of communication was very similar. One thing

Looking at the histogram of mental demand results from round two of testing shows that, with the addition of visual communication along with the audio, there is a shift in the mental demand results. Similar to the time results, there is a right skew in the mental demand. The HoloLens has more of a shift than the projector and the screen.

It is important to note that even though there are not too many differences in the average mental demand, the shift in the overall distribution tells us that the visual communication along with the audio is reducing the test subjects' mental demand. We expect that if the difficulty of the task were increased, we would see even more of a shift in the data towards lower extreme TLX. In addition, the data shows that the visual communications all have interquartile ranges that are closer together when compared to audio communication. The TLX results demonstrate that we are able to combine audio communication with visual communication without the visual communication significantly changing the average TLX.

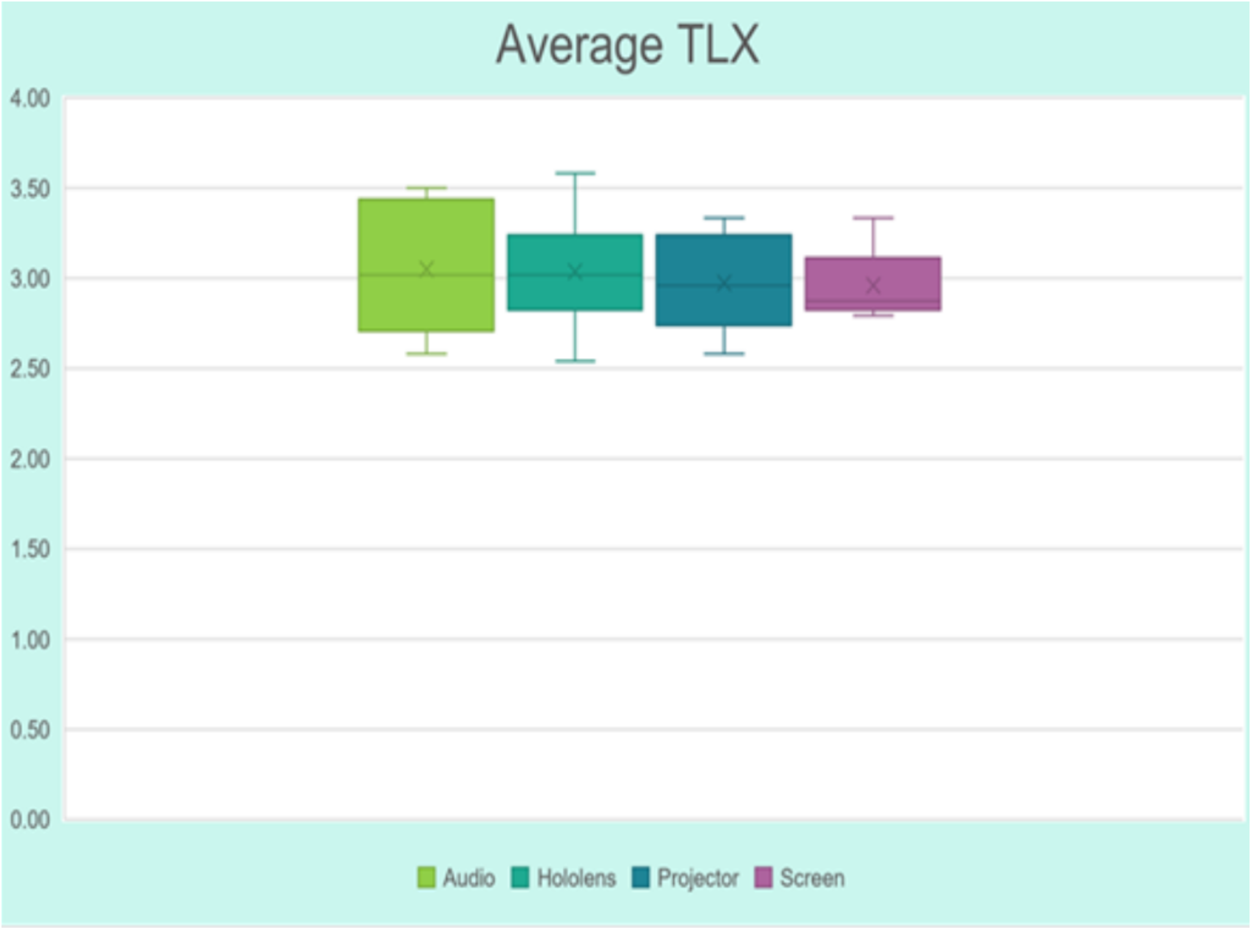
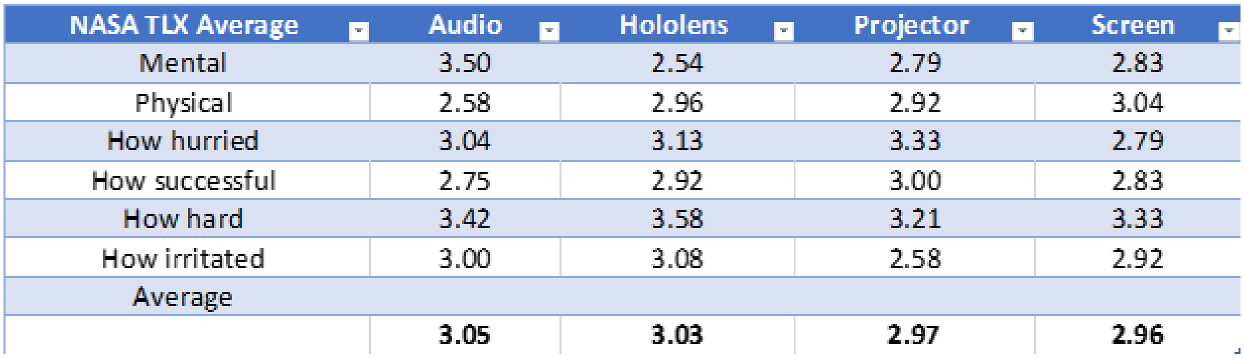


Figure 42: NASA-TLX average scores for Round 2 testing.



NASA TLX Average	Audio	Hololens	Projector	Screen
Mental	3.50	2.54	2.79	2.83
Physical	2.58	2.96	2.92	3.04
How hurried	3.04	3.13	3.33	2.79
How successful	2.75	2.92	3.00	2.83
How hard	3.42	3.58	3.21	3.33
How irritated	3.00	3.08	2.58	2.92
Average				
	3.05	**3.03**	**2.97**	**2.96**

Figure 43: NASA-TLX average scores for Round 2 testing (tabulated).

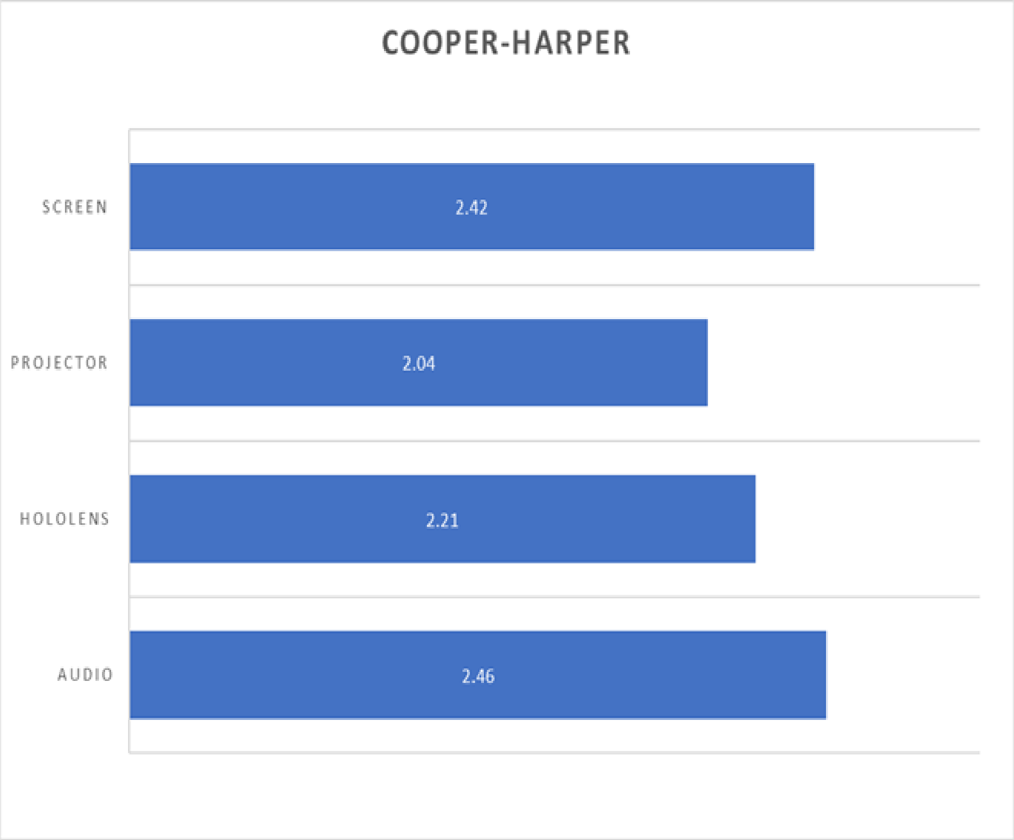
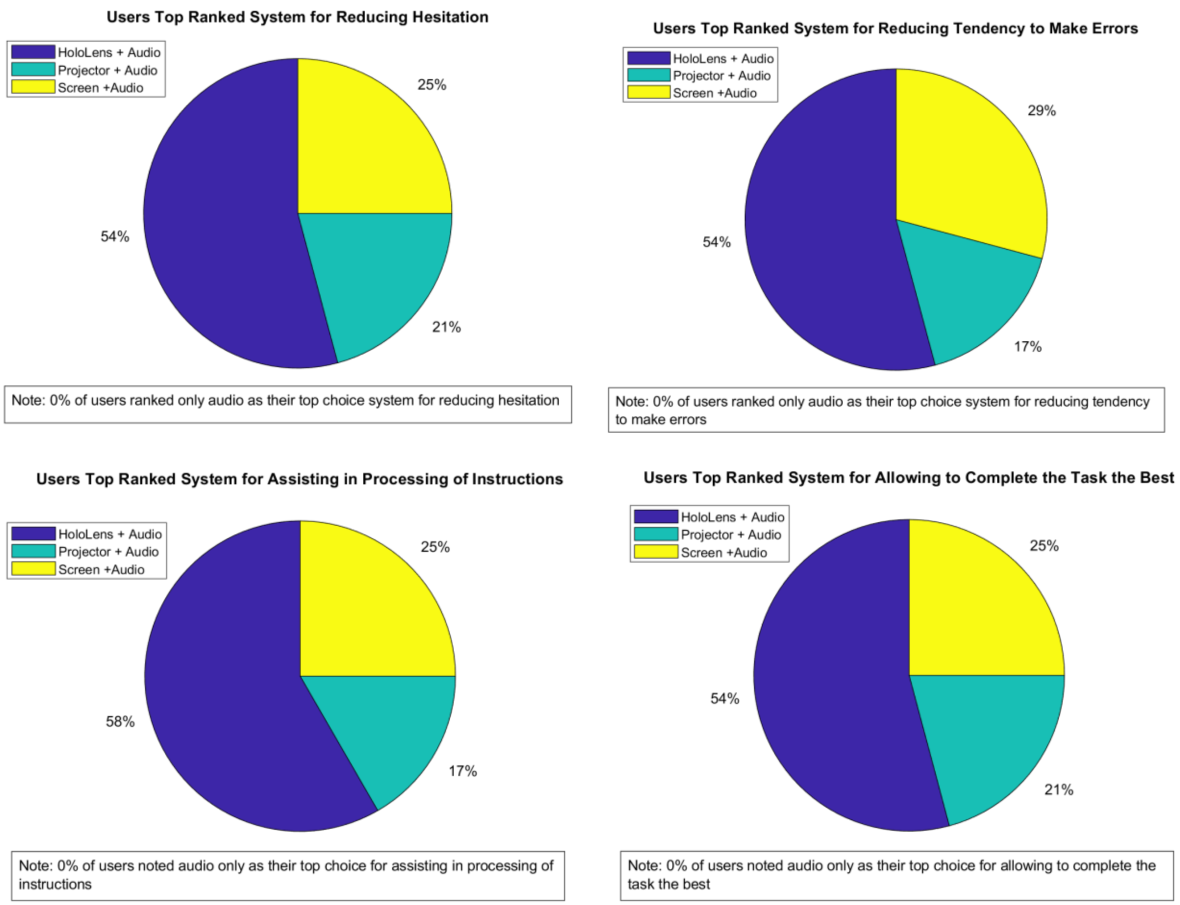


Figure 44: Average Cooper-Harper ratings for each type of test.

3.5.1.3 User Preferences (Rankings)

After testing with each system, the user was asked to rank the four systems (from 1st-4th) on which system they found best for each of the following: reducing hesitation, reducing tendency to make errors, assisting in processing of instructions, and allowing to complete the task the best. The pie charts below in **Figure 45** show the portion of systems that users ranked as their 1st choice for each category.


Users Top Ranked System for Reducing Hesitation

Note: 0% of users ranked only audio as their top choice system for reducing hesitation

Users Top Ranked System for Reducing Tendency to Make Errors

Note: 0% of users ranked only audio as their top choice system for reducing tendency to make errors

Users Top Ranked System for Assisting in Processing of Instructions

Note: 0% of users noted audio only as their top choice for assisting in processing of instructions

Users Top Ranked System for Allowing to Complete the Task the Best

Note: 0% of users noted audio only as their top choice for allowing to complete the task the best

Figure 45: Users' top choice rankings.

As seen from the charts, none of the users chose audio only as their top choice for any of the questions asked. This indicates that every user that tested preferred to have some sort of visual aid to complete the task. Additionally, each question shows a similar preference of system plus/minus one person shifting their response. The majority of users (about just over half) chose the HoloLens and audio as their top preference for all four questions. About a quarter of users ranked the external screen display and audio as their top preference. Lastly, just under one quarter of users ranked our HUD and audio as their top choice. In conclusion, this indicates that users preferred the HoloLens and audio as their top choice and audio only as their last choice.

3.5.1.4 Confounding Variables

There are a few confounding variables that could have influenced the data. These variables include the learning curve, familiarity with AR technology, and comfort within the helmet. In order to reduce the effects of decreasing time across tests from learning the tasks, the order of the systems tested was randomly generated for each user. However, there is still a learning curve; the tasks for each test were similar, it is likely the users found more efficient ways to complete certain tasks by their third or fourth time doing them.

Another confounding variable is familiarity across users with technology similar to the HoloLens. Some users were likely more comfortable with the HoloLens if they have had experience wearing headsets in gaming or have had experience with other AR technologies in general. On the other hand, to some users the display from the HoloLens was disorienting and unfamiliar, which could have led to a higher completion time due to adjusting to something new.

With reference to user rankings, it is inevitable that the HoloLens is a well-known and robust system. The prototypes of our helmet-mounted display and external screen display had simple designs and functioned only to provide a display for testing, due to constraints in resources. Even though the HoloLens was only used to show the same display, users may have ranked the HoloLens higher due to the high-level technology and capabilities of the system.

The helmet comfort could have also been a confounding variable. Users with a larger frame and shoulders tended to have no issues with helmet stability. However, some users with a smaller frame found the helmet to be unstable and hard to support. Frequently, these users moved slower in order to ensure the helmet would not fall off. In result, this variable could have led to a discrepancy in completion times.

3.5.1.5 Overall Results Discussion

As discussed, results from testing were collected to demonstrate efficiency, comfort, and user preference across four different methods: audio only, audio and HoloLens, audio and helmet-mounted display, and audio and external screen display. It was concluded that the audio and helmet-mounted display showed a statistical significance in difference in time taken to complete the task in comparison to the control, audio only. The TLX showed a similar mean for all four systems, meaning that the average of the users' inputs indicated similar comfort for all systems tested. This was interesting because in an EVA space application, the HoloLens would likely not be comfortable as it takes up additional space inside the helmet and cannot be adjusted once the EVA has started. Another interesting note can be seen in the discrepancies between the results from efficiency and user preference. Slightly over half of users ranked the audio and HoloLens as their top choice for the four questions discussed in the User Preferences section. The other half of users preferred either the helmet-mounted display or external screen display. Across the visual systems, users' preferences differed from how they performed in terms of time they took to complete the task. Our helmet-mounted display showed highest efficiency, but the HoloLens display was ranked higher by users. From the data statistics, it is likely that the audio and HoloLens system did not show a statistical significance in efficiency due to the high variance across users.

In conclusion, the results show that users perform better and prefer to have some sort of visual display in completing the tasks. Although the majority showed a preference for the HoloLens, the results show that the other visual displays help the user perform just as well and with equal comfort.

3.5.2 Discussion (Non-data Observations)

There were several observations that we made throughout the course of our testing which have greatly impacted the conclusions we made from our data. One of the more important factors we observed for each display method we used was the preparation time required to activate each of them. For our prototypes, the preparations were relatively brief, as we had to simply power on the device and connect our HDMI receivers to ensure the signals were being properly transmitted. This process took only several minutes, and if this technology were fully implemented into a spacesuit the preparation time would likely shorten even further as there would not have to be a transition between the two display methods.

Another factor that was worth considering was the possibility of requiring different sizes and configurations of each display in order to accommodate the needs of different people. In the case of the projection setup, some of the subjects had difficulty with the placement of the projector and the display screen in our HUD, as their head or hair would block off part of the projection at certain angles. As both the projector and projection screen setups were fixed in the helmet, this was an unavoidable difficulty which likely had an effect on the data for the projection test. The external screen display prototype also had issues with how it was placed and angled relative to each subject's field of view, but as the screen itself was attached to an adjustable phone holder this issue was more negligible. Both of these experiences do show that a display system should be adjustable to fit with each user's dimensions and preferences in order to minimize discomfort and optimize the flow of information. This is, of course, easier for a screen system than a projector system, as with the projector system you have a much more limited set of angles to work with in terms of ensuring parts of the display are not cut out and are easily visible.

The HoloLens, on the other hand, had several issues in terms of preparation and usage. While the helmet model that we used for our experiments was able to fit the headset inside of it, in order to do so we were required to disassemble the helmet into its top and bottom halves before inserting or removing the HoloLens. This process, while not complicated, was extremely tedious, and the added difficulty of setting up the HoloLens program correctly added a significant amount of time to the setup of this phase in testing. In addition, the HoloLens had an irritating tendency to slide off of the test subject's face, which in some cases required us to stop and restart the test in order to take the helmet apart once more and fix the problem. In other cases, the participant was forced to carry on using only the audio communications. Neither of these factors make the HoloLens a very appealing option in practical use, as an astronaut on an EVA often does not have the luxury of being able to start again if the headset slips off. Instead, the head-mounted display becomes a liability and a risk to the astronaut's safety.

There were also some overall observations which we made regarding our overall testing procedure. While we did apply a randomization factor to each individual procedure for each run-through of the test, we were not able to introduce complete randomization into the code. Instead, we were forced to have several of the first and last steps stay the same for all tests, which made it much easier for the participants to recognize and learn the procedure over the course of all four trials. As a result, the learning curve of the participants had a greater impact on our tests than was preferable. However, we mitigated the issue to the best of our ability by randomizing the order in which the tests were performed. By having each communication method be first and last in the order of our testing, we were able to compensate for the decreased time and increased user confidence caused by the learning curve of the participants.

4 Conclusions and Future Work

4.1 Future Directions for Project

After developing and testing two different approaches to a helmet-mounted visual aid alongside audio communication and the HoloLens, there were a few clearly some changes that could be made to the design to improve user-friendliness. After round two of testing, participants tended to score the HoloLens highest on ease-of-use, primarily because the display was right in front of them. Our helmet-mounted display and external screen display prototypes both required the wearer to take their eyes off of the task they were performing, which some participants found detrimental to their performance. A newer iteration of the in-helmet prototype should move the display directly in front of the wearer to remedy this. Another change would be to improve the resolution and contrast of the display, since some participants remarked that the built-in projection system and screen were hard to read and occasionally displayed colors incorrectly. A next-generation prototype could make use of more purpose-built optics and hardware to produce a clearer display, taking note of the difficulties faced when attempting to do just that when developing version one of this project’s prototype.

Another next step for this project would be to build a more realistic astronaut helmet with the display system inside. This would include making the helmet airtight so it could be tested in the neutral buoyancy tank (at the Space Systems Lab, UMD) as well as integrating an actual audio communication system. Since participants still generally relied on audio communication to some extent even when a display was present, the next prototype should continue to enable audio communication. For the tests performed, direct verbal communication without microphones or earpieces was sufficient, but having a microphone and headset in the helmet would allow for a wider range of tests where the participant isn’t sitting still or right next to the tester. Waterproofing the electronics onboard the helmet would also be necessary for it to operate in the neutral buoyancy tank. An extension of this would be to make the display system within the helmet fully self-contained, so that test participants could operate long distances away from testers and still receive both audio and visual instructions. Building in a small onboard computer such as a raspberry pi to handle instruction delivery could accomplish this and was briefly explored in this project before being deemed too redundant for the tests performed.

4.2 Our Systems and Space Applications

Both the external screen display and the heads-up projector display systems will need to go through further design iterations and testing before either could be spaceflight qualified. Most notably, the transition from 3D printed prototype helmet with integrated projector system to a spaceflight qualified helmet with an integrated HUD would present hazards given operation in the oxygen environment within the suit. One possible way to avoid having electronics operate within the pressurized portion of the helmet would be to reroute the optical path such that the display serves as one module which can be attached and removed from the top of a standard EVA helmet. This would likely require further bench testing similar to that described previously; however, it would also allow for the ability to implement design changes to ensure that the controls on the projector were easily accessible, and usable, for the astronaut.

Aside from issues with the HUD operating in the pressurized suit, both displays will also need to be adjusted for use in the varying and harsh thermal environment of space. Furthermore, the long duration of continued use during extravehicular activity also presents thermal as well as power complications. These issues will likely require methods of insulation and/or cooling to be explored in addition to an overall analysis and reconfiguration of the current power system used for short testing periods on Earth.

From a software perspective, it is easy to envision transitioning from displaying the simple task instructions and images to displaying information needed to assist the astronauts during the EVA. Additionally, linking data for their vitals and metrics regarding things like suit pressure, oxygen, and battery life could be implemented if found to be necessary or helpful.

Finally, the large variation in the lighting conditions of space provide further challenges. No testing has been done on the prototypes in unfavorable lighting conditions on Earth. These would be conditions where the teleprompter glass of the projection helmet-mounted display was backlit, or where the screen system was brightly illuminated from the front. In order to ensure proper functionality and readability in space, more testing will need to be done in these areas. The results of those tests might suggest that a method of automatically adjusting the brightness of the displays is needed.

While the prototype designs developed for the tests conducted and outlined in this paper proved to be functional in a lab setting on Earth, there are many more challenges that will need to be overcome before the design can be scaled for space applications. A great deal of testing and iterative design will be required.

5 Works Cited

1. M. Garcia, “Space Station Spacewalks,” *NASA*, 30-Jul-2015. [Online]. Available: http://www.nasa.gov/mission_pages/station/spacewalks.
2. “Moving around Mars.” NASA Mars Exploration Rovers. https://mars.nasa.gov/mer/mission/timeline/surfaceops/navigation/ (accessed Feb. 13, 2020).
3. “Hololens.” Microsoft. https://www.microsoft.com/en-us/hololens (accessed Mar. 30, 2020).

4. “Heads up display for spacesuit.” Manas. https://manas.tech/projects/aouda/ (accessed Mar. 30, 2020).
5. G. Wetzstein, “Computational Near-Eye Displays: Engineering the Interface to the Digital World,” The National Academy of Engineering, vol. 46, no. 4, pp.5–9, Winter 2016.
6. C. E. Carr, S. J. Schwartz, and I. Rosenberg, “A wearable computer for support of astronaut extravehicular activity,” in the 6th International Symposium on Wearable Computers, Seattle, WA, 2002.
7. NASA OIG illustration of data from NASA and Thomas, Kenneth S., and Harold J. McMann, U. S. Spacesuits, New York, NY: Springer New York, 2012.
8. NASA Office of Inspector General, “NASA’s Management and Development of Spacesuits.” *NASA*, 26-Apr-2017. [Online]. Available: https://oig.nasa.gov/docs/IG-17-018.pdf.
9. “Spacesuit Basics.” Available: https://www.nasa.gov/feature/spacewalk-spacesuit-basics
10. E. Mahoney, “A Next Generation Spacesuit for the Artemis Generation of Astronauts,” *NASA*, 28-Oct-2019. [Online]. Available: http://www.nasa.gov/feature/a-next-generation-spacesuit-for-the-artemis-generation-of-astronauts.
11. “Xemu Development Unit 4.” Available: https://www.nasa.gov/image-feature/exploration-emu-xemu-development-unit-4/
12. N. Patrick, J. Kosmo, J. Locke, L. Trevino, and R. Trevino, “Extravehicular Activity Operations and Advancements.” NASA. [Online]. Available: https://www.nasa.gov/centers/johnson/pdf/584725main_Wings-ch3d-pgs110-129.pdf

13. M. Dunn, “Astronaut completes spacewalk without helmet camera, lights,” *ABC News*. [Online]. Available: https://abcnews.go.com/Technology/wireStory/nasa-conducts-female-spacewalk-battery-work-68295129.

14. E. Howell, “Astronauts complete 4-spacewalk marathon to fix space station’s $2 billion antimatter detector,” *Space.com*. [Online]. Available: https://www.space.com/spacewalking-astronauts-repair-space-station-antimatter-detector.html.

15. O. Doule and L. Poulet, “Ergonomy of Head-Mounted Displays Inside Analog Spacesuit-Mars Analog Extravehicular Activities,” in AIAA SPACE 2014 Conference and Exposition, Anaheim: AIAA, 2014, pp. 1-26.

16. R. Jose, G. A. Lee, and M. Billinghurst, “A Comparative Study of Simulated Augmented Reality Displays for Vehicle Navigation,” in Proceedings of the 28th Australian Conference on Human-Computer Interaction, New York City: OzCHI, 2016, pp. 40-48.

17. V. Vechev, A. A. Unl¨uer, and A. Dancu, ”Medium and Transition in Mobile ¨ Projected Interfaces,” in Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, 1st ed. New York: ACM, 2015, pp. 1134-1137.

18. K. Ito, H. Nishimura, and T. Ogi “Head-up Display for Motorcycle Navigation,” SIGGRAPH Asia 2015 Heads-Up Displays and Their Applications, SIGGRAPH, New York City: 2015, ch. 10, pp. 1-3.

19. B. C. Kress and W. J. Cummings, “Towards the ultimate mixed reality experience: HoloLens display architecture choices” in Society for International Display, vol. 48, no. 1, pp. 127-301, May 2017.

20. J. Blattgerste et al., “Comparing conventional and augmented reality instructions for manual assembly tasks” in Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments, Island of Rhodes, Greece: Jun 2017, pp. 75-82.
21. M. Garon et al., “Real-time high resolution 3D data on the HoloLens” in 2016 International Symposium on Mixed and Augmented Reality Adjunct Proceedings, Yucatan, Mexico: Sep 2016, pp. 189-191.
22. H. Murat, H. De Smet, and D. Cuypers, “Compact LED Projector with Tapered Light Pipes for Moderate Light Output Applications,” in Research Gate, 1st ed. Ghent, Belgium: TFCG, 2006, pp. 117-123.
23. S. Tachi, “Telexistence and Retro-reflective Projection Technology (RPT),” in The Fifth Virtual Reality International Conference, France: VRIC, 2003, pp. 1-69.
24. F. A. Burns, “Adaptable HUD Mount,” in US6072444A, 1st ed. Alexandria: USPTO, 2000. 36
25. C. Wei Hsu et al., “Transparent Displays Enabled by Resonant Nanoparticle Scattering,” in Macmillan Publishers Limited, 5th ed. London, United Kingdom: Nature, 2014, pp. 6-7.
26. L. Zhou et al., “All-Organic Active Matrix Flexible Display,” in Applied Physics Letters, 88th ed. Melville: APL, 2006, pp. 1-4.
27. A. Sugimoto et al., “Flexible OLED Displays Using Plastic Substrates,” in IEEE Journal of Selected Topics in Quantum Electronics, 10th ed., Piscataway: 2004, pp. 107-114.

28. P. K. Jain et al., “Calculated Absorption and Scattering Properties of Gold Nanoparticles of Different Size, Shape and Composition:,” in Journal of Physical Chemistry B, 110th ed. Washington, D.C.: J. Phys. Chem. B, 2006, pp. 7238- 7248.

29. Human Integration Design Handbook, 1st ed., NASA, Washington, D.C., 2015, pp. 13-1294.

30. M. J. Lewis and D. J. Schuck, “Results of a helmet-mounted display precursor demonstration at desert RATS,” in 42nd International Conference on Environmental Systems, San Diego, CA, pp. 1-33.

31. S. P. E. Russell, “The extravehicular maneuvering unit’s new long life battery and lithium ion battery charger,” in AIAA Space 2010 Conference and Exposition, Anaheim, CA, 2010.

32. S. G. Hart, “NASA Task-Load Index (TLX)”, NASA Ames Research Center, Moffett Field, CA, USA, Jan. 1986.

33. D. G. Mitchell, “Fifty Years of the Cooper-Harper Scale,” presented at the AIAA SciTech Forum, San Diego, CA, Jan. 2019.

34. B. Donmez, A.S. Brzezinski, H. Graham, and M. L. Cummings, “Modified Cooper Harper Scales for Assessing Unmanned Vehicle Displays,” MIT Humans and Automation Laboratory, Boston, MA, USA, HAL2008-06, Dec. 2008.

35. A.M. Braley, B. Nuernberger, and S. Young Kim, “Augmented Reality Improves Procedural Work on an International Space Station Science Instrument,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 61, No. 6, pp. 866-878, Jan. 2019.

36. Team VISOR, “Development of a Heads-Up Display for Extravehicular Activities,” ICES, 2019.
37. V. Thakur and J. Ferri, “TI DLP Pico Technology for Aftermarket Head-up Displays.” Dallas, Sep-2016.
38. S. G. Kwak and J. H. Kim, “Central limit theorem: the cornerstone of modern statistics,” Korean journal of anesthesiology, Apr-2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5370305/. [Accessed: 30-Mar-2020].
39. SPSS. [Online]. Available: https://psychology.illinoisstate.edu/jccutti/138web/spss/spss3.html. [Accessed: 30-Mar-2020].

6 Bibliography

1. H.-Y. Kim, “Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis,” Restorative dentistry & endodontics, Feb-2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3591587/. [Accessed: 30-Mar-2020].
2. “Kurtosis: Definition, Leptokurtic, Platykurtic,” Statistics How To. [Online]. Available: https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/statistics-definitions/kurtosis-leptokurtic-platykurtic/. [Accessed: 30-Mar-2020].
3. Analytic Hierarchy Process, AHP. [Online]. Available: https://scbuk.com/ahp.html. [Accessed: 30-Mar-2020].