Development of a Heads-Up Display for Extravehicular Activities

Katheryn Fox, Radhika Karsalia, Jillian Kunze, Christoph Neisess, Zachary Peters, Roshan Rao, Brady Sack, Matthew Sieh, Ryan Skoletsky, Shelly Szanto, Matthew Wilkin[1], and David Akin[2]
University of Maryland, College Park, Maryland, 20742, USA

Spacesuits represent a significant burden on the wearer in terms of isolation and lack of dexterity. Compared to the current population, never out of reach of a smart phone for access to myriad databases, the spacesuited astronaut is limited to their knowledge, training, and verbal guidance from mission control. These capabilities will be increasingly ineffective in future space exploration missions, where time delays or lack of line-of-sight to Earth will effectively eliminate mission control as a factor, and neither knowledge nor training will be wholly adequate to support the crew through many hundreds of hours of EVA on an extended mission to the moon or Mars. To address this problem, the University of Maryland has been performing a multi-year investigation of options to enhance data presentation to spacesuited crew, through the development of helmet-mounted non-obstructive visual overlays. Initial efforts focused on the provision of head-mounted displays for one or both eyes, remotely controllable to pivot up onto the forehead when not in use. This proved cumbersome, and two current efforts are proceeding in parallel. In the first, an oversized helmet was developed for a UMd MX-D spacesuit simulator which allows the wearing of a Microsoft HoloLens inside the helmet. This allows the presentation of a high-resolution visual overlay which does not obstruct the actual visual field, but the HoloLens has a very small field of view. The parallel effort is focused on a helmet specifically designed to incorporate an internal projector and optics to provide an overlay on an external visor assembly, which can be raised or lowered as desired. Details of both systems are presented, along with results from both laboratory and initial field testing. The paper includes lessons learned and plans for further development activities.

Nomenclature

AR	=	Augmented Reality
EMU	=	Extravehicular Mobility Units
EVA	=	Extravehicular Activity
HUD	=	Heads-Up Display
HMD	=	Helmet Mounted Display
LCD	=	Liquid Crystal Display
MR	=	Mixed Reality
SSL	=	Space Systems Laboratory
TFT	=	Thin Film Transistor
UMd	=	University of Maryland

[1] Undergraduate Student, University of Maryland
[2] Director, Space Systems Lab; Associate Professor of Aerospace Engineering

I. Introduction

As the space industry looks forward to human exploration of Mars and eventually beyond, astronaut self-sufficiency becomes of paramount concern. The further humans venture into the solar system, the more independent astronauts must become as communication with those on Earth becomes infeasible, and neither knowledge or training will be wholly adequate to support the crew through the many hundreds of hours of EVA that are required on extended missions.

For independent exploration to become a reality, the current system for communicating critical information to astronauts during an EVA requires extensive redesign. NASA Extravehicular Mobility Units (EMU), originally designed in 1974 and still used to this day, currently leave the spacesuited astronaut exclusively limited to knowledge, training, and explicit verbal guidance from mission control on Earth [1]. With real-time communication entirely infeasible for missions to Mars, where communications delays average around 13 minutes, it becomes necessary for astronauts to have access to procedural information and suit vitals through a different medium. For this reason, the development of heads-up displays (HUD) for future space suit helmets is necessary for the advancement of human space exploration.

The research presented in this paper is divided into two parallel efforts: (1) the development of a high resolution HUD on the Microsoft HoloLens for use in an oversized helmet and (2) the design of a helmet mounted display (HMD) which uses a projector. Human factors testing was conducted with the HoloLens HUD, and the results from these tests will be used to inform the development of the helmet-mounted system.

II. Development of HoloLens HUD

The Microsoft HoloLens, a mixed reality (MR) headset, was chosen as a development platform due to its high resolution and well-supported developer environment. These assets permitted us to quickly design and test a simple HUD (the design of which is detailed below) on Unity, a cross-platform game engine that supports the HoloLens.

The final augmented reality (AR) display, shown below as a developmental version (Figure 1), consists of two main features to assist test participants in completing each of our test runs in a timely fashion, which is discussed in depth in the next section. The display features an enlarged written statement of the instructions along with a scaled-down color image of the corresponding task on the board. These are in the user's upper peripheral vision at a farther domain than the task board at approximately 1-2 meters; this corresponds to a clipping plane in the Unity platform of 0.85. This clipping plane value was chosen based on Microsoft's recommended user settings to avoid optical strain. By placing the text display and image in the upper peripheral, we avoid the intersection of the augmented images with direct line of sight to the task board, which lies in a domain approximately 0.30 meters from the user's eyes. The text is displayed in a green wavelength, since the human eye is best adapted for detecting wavelengths around 507 nm [3]. The user controls the transitioning of the text and images through the HoloLens voice input feature by stating the words "next" or "back". With the latest HoloLens update, both the text and image track with user's head movements automatically to avoid lag and to provide the test subject with a continuous and consistent display throughout the tasks. In the future, we will also look into interpupillary distance calibration to improve the user's unique experience. The display was tested in a room with standard lighting, minimal to no objects in the surrounding environment, and a black backdrop surrounding the test board. The display notifies the user through text when all successful tasks have been completed.

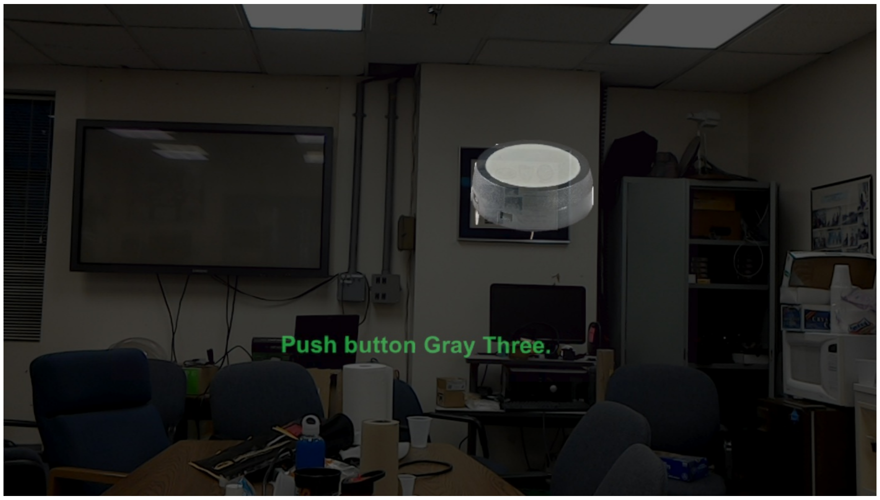


Figure 1. An example instruction projected through the HoloLens.

III. Testing of HoloLens HUD

A. Testing Methodology

Initial experimentation consisted of three tests in which the participant completed simple tasks on a task board, shown below (Figure 2). Each test featured a different delivery medium for task instructions: (1) audio instructions, (2) AR display on the Microsoft HoloLens, and (3) both audio and HoloLens instructions. Time taken to complete each test was recorded for each participant, and after each test the participants were asked to fill out a short survey, which included NASA Task Load Index (TLX) evaluations and Cooper-Harper ratings. A detailed discussion of the data is found in the next section.

The pool of subjects was a random selection of thirty adults who were recruited using physical advertisement and online recruitment before being screened with a survey prior to testing. The requirements for individuals to qualify as test participants were that they must have either 20/20 or corrected vision, have normal hearing, and not be colorblind. If, on the survey, a potential participant marked that they were not sure of their vision, a standard eye exam and colorblind test were used in order to determine their eligibility. We also asked participants their height to see, in the future of this project, if taller- or shorter-than-average people might have more or less difficulty with the tests.

The task board was organized to loosely resemble a control panel or panel that requires maintenance 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. In addition, it 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 four different carabiners inside. The task board was designed to mimic simple, generalized tasks that participants could complete without prior knowledge or training. The task board was connected to a structure made of 80/20 aluminum extrusion that angles it towards the test participants.

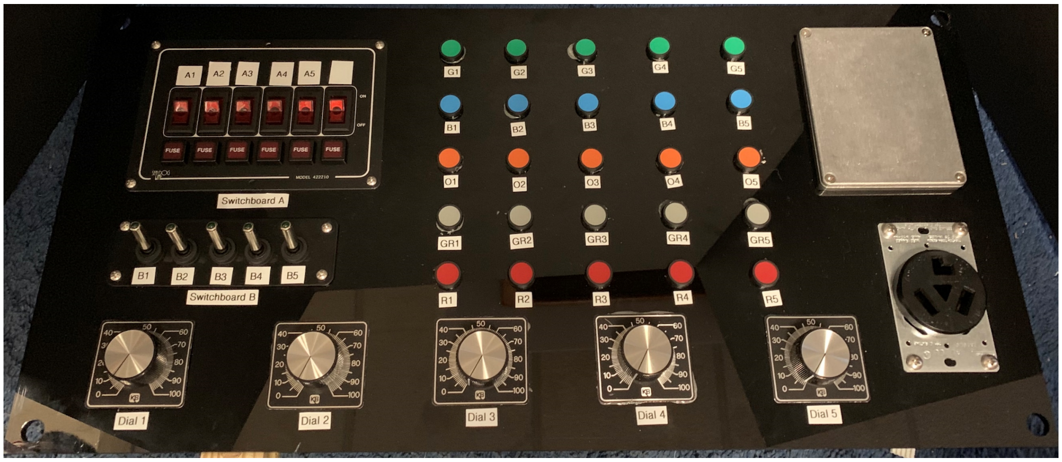


Figure 2. The layout of the "control panel" board which participants used during testing.

In each of the three tests, participants ran 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 task board before returning them to the box and screwing the lid back on. While the steps to open and close the metal box and install and remove the carabiners were deliberately made the very first and last steps, every other one of the 50 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. Specifically, audio communication involved direct person-to-person interaction with a specific researcher, where one of our team members would read off the directions to the test subjects. The subjects were able to ask their contact 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 proceed through the procedure, and was able to view previous steps as desired. 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 communication with the team member reading the procedure aloud. Participants would use the voice command "Next" to advance to the next step or say "Back" if they needed to repeat the previous step.

B. Initial Results

i. Task Performance

The analysis was broken into two major portions. In the first set of analysis 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 at a 95 percent confidence interval. To further determine the relative differences, we conducted a paired t-test for difference in means at a 95 percent confidence interval. The first set of analyses tested for statistical significance in performance time to determine if one treatment had a negative impact on performance speed. Our initial results from the 30 test subjects indicates that there is a statistically significant difference between users as well as between the three forms of communication that users were given to perform the simulated EVA task. The paired t-test for difference in means indicates that, regarding time, there is a strong statistical difference in mean performance time between the audio communication method and the combined audio and HoloLens communication method, which leads to the conclusion that there is a slower task performance when audio is used without a spatial aid. Although smaller, there is 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 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. Refer to the Appendix section for a view of the raw data and more detailed analysis.

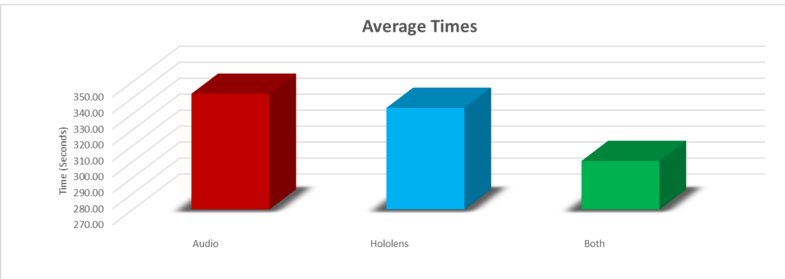


Figure 3. Bar Graph of Average times n = 30

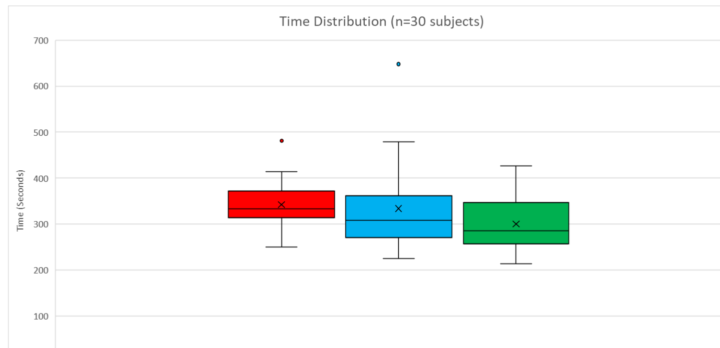


Figure 4. Box and Whisker Plot of times n = 30

ii. TLX and Cooper Harper

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 ratings, we once again conducted a two-factor analysis of variance and blocked for user variation at a 95 percent confidence level. For both the NASA TLX and the Cooper-Harper there was no statistically significant difference 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. 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, despite briefings and clearly-readable labels, 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. For a view of the raw data, again refer to the Appendix section.

iii. Statistical Analysis

Our initial results from the 30 test subjects indicate that there is a statistically significant difference between users as well as between the three forms of communication that users were given to perform the simulated EVA task. The paired T-test for difference in means indicates that there is strong statistical difference in mean performance time between the audio communication and both audio and HoloLens communication, which indicates 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 and combined HoloLens and audio. Finally, there is no statistical difference in the mean performance between audio 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 among the 30 test subjects shows 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, and that with proper training the HoloLens may improve performance even further.

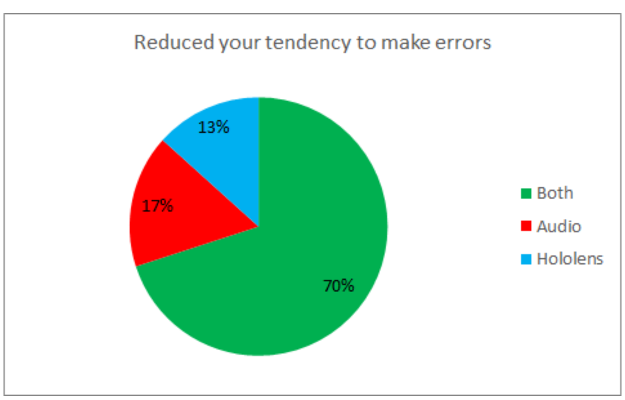
A second round of analysis was also conducted on the NASA TLX scores and the Cooper-Harper ratings in order to see if the three different forms of communication had statistically significant differences, in order to determine the overall effect on user comfort. For both the NASA TLX and the Cooper-Harper ratings there was no statistically significant difference between the three user interfaces at a 95 percent confidence interval; however, the physical demand and mental demand portions of the NASA TLX came close to indicating a significant difference. Breaking the data down further and analyzing the paired T-test for difference in mean TLX value displayed higher levels of mental demand for HoloLens or audio alone when compared to both interfaces combined. 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 audio communication. Combined with the relative difference in mental demand, this helps indicates that the HoloLens is improving performance time without introducing any additional significant user strain at a 95 percent confidence interval.

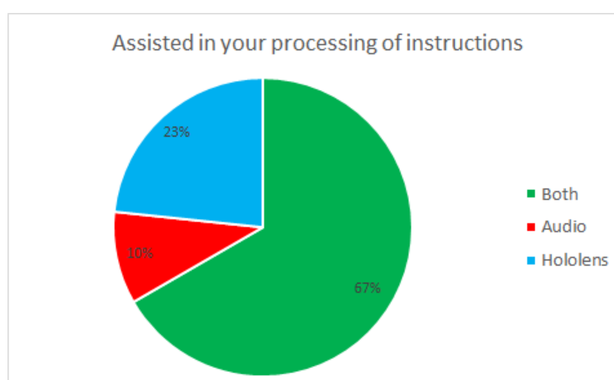
iv. Survey

Test subjects were asked to complete a survey where they would rank the three forms of communication based on four different metrics: (1) reduced hesitation to complete the task (i.e. increased confidence) (2) assisted in processing of instructions (3) reduced tendency to make errors and (4) allowed user to complete the task the best in general. All thirty test subjects were asked these four ranking questions one time after all three tests were completed.

The majority of test subjects ranked both HoloLens and audio combined as the best method in these four metrics. The overall results were recorded in the figures below. Out of all of the test subjects, 64% said that this method reduced hesitation to complete the task, 70% said that it reduced their tendency to make errors, 67% said that it assisted in the processing of instructions, and 63% said that it in general allowed them to complete the task the best. The HoloLens-only method was the second most highly ranked, except in the category "Reduced tendency to make errors."

We also asked test subjects to describe how much they paid attention to the HoloLens instructions compared to the audio instructions during the combined HoloLens and audio test. The majority responded that they paid more attention to the HoloLens than they did to the audio instructions. Some stated that they paid equal attention to both, with one person reporting that they primarily relied upon the HoloLens 75% of the time and the audio 25% of the time. A few subjects reported that they used the visual pictures to decide what to do next but waited for the audio to be finished before advancing to the next instruction in order to confirm they completed the task correctly. One subject reported that, for the simpler tasks, they were more dependent on the visual instructions. In general, the feedback stated that the HoloLens instructions gave them an advantage as a way to confirm that they were hearing the audio instruction correctly. As an example of this, one of the tasks is to push a blue button labeled "B1" but there is also a





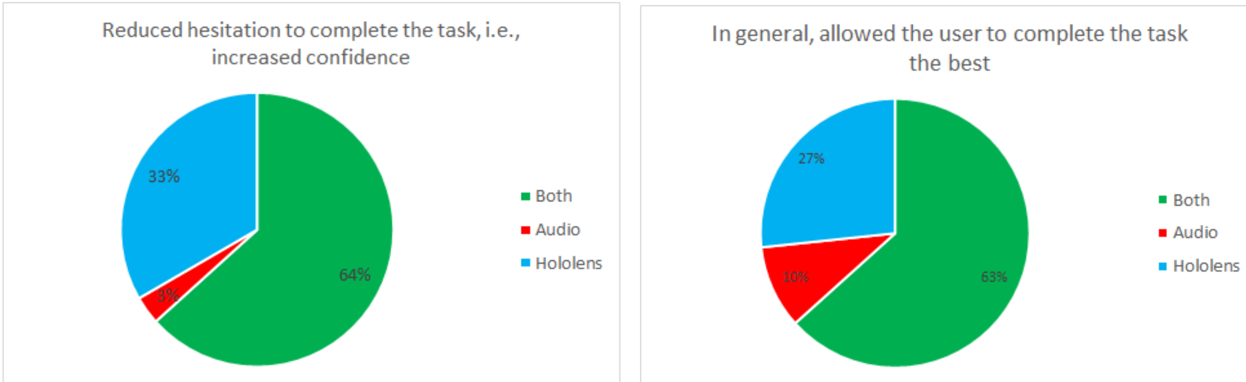


Figure 1. Survey results n = 30

switch on switchboard B called "B1." When given this audio instruction, users would often flip the switch when they were supposed to press the button, and vice versa. When given the HoloLens' image of the button or the switch they could clearly see which one they were supposed to do and corrected that mistake.

v. A comment on learning curve in Human Factors testing

For all of human factors testing there is the factor of the learning curve to consider. Our test subjects were given a random order in which they would complete the three tests. The figures below display the distribution of the test orders that our subjects underwent. We hypothesized that there would be an added learning curve with the HoloLens and HoloLens and audio combined tests because, for 97% of our test subjects, this was their first time using augmented reality or wearing the HoloLens. There was a clear factor of learning with each test. The first test that subjects would do would usually take longer due to the subject's looking around the board to find the specific task that they were told to complete. By the second and third tests, the subjects were comfortable with the locations of the tasks on the board and able to complete each task with more ease. During their third test, a few subjects even

Figure 2. Distribution of order of tests n = 30

predicted and began performing the last step (screwing the box lid back on), as it and the previous step (putting the carabiners back into the box) were always the last steps for each test, before they were told to do so.

IV. Development of HMD

All of our preliminary testing with the HoloLens is to inform the development of our own prototype Heads-Up Display, which will be incorporated into the existing schema of an astronaut's helmet. While the HoloLens provides an excellent platform for development and testing, it has several drawbacks that disqualify it from use in real EVAs. The field of view is prohibitively narrow, subtending only 35 degrees of view, which quickly becomes frustrating for the users. The device is also bulky and uncomfortable when worn for extended periods of time, and, since it is mounted to the head, can accelerate user fatigue. Along similar lines, the user cannot adjust the HoloLens once inside the helmet during an EVA. Finally, the headset is too large for a standard NASA helmet and requires an oversized helmet for testing. Therefore, our team set out to develop a heads-up display that has both a wider field of view and is more comfortable for the user.

To design our HUD, we have been exploring potential layouts for the screens, lenses, mounts, and electronics within the helmet. Our current plan is to have two projector assemblies attached to the upper sides of the inside of the helmet via a ball and socket mount. However, specifications on readability and image size from the HoloLens testing will need to be incorporated into the development of our first prototypes, as well as constraints imposed by the size of the helmet and the short projection distance.

A. Physical Design

The overall helmet and HUD design consists of many parts. The helmet that we are using is the MX-D spacesuit simulator helmet, which is a 16 inch diameter spherical helmet 3D printed in PLA plastic, with a rotating openable visor. This unit was specifically designed to accommodate the Microsoft HoloLens head-mounted display system, and is large enough to accommodate the hardware we wish to implement. We have added the hardware components for the HUD system, which include two screens, lenses, a mount for the screens, and electronics to feed the information from a computer to the screens. The two screens are Liquid Crystal Display (LCD) screens mounted to each side of the helmet and angled in such a way to reflect the images displayed into the eyes of the astronaut. We expect to use LCD screens with a 2.2 inch diagonal. However, while smaller screen sizes may be explored, including anything larger is not necessary for the design and would limit free space in the helmet. The two mounts are located on either side of the astronaut's head and will be discussed in more detail below.

The helmet will also include an electrical system to transport the task information to the screens, consisting of a Raspberry Pi single-board computer, the LCD screens, a fan, and connecting wires. These wires will be run from the Pi, through the back of the helmet, and plugged into the LCD screens. The software system to generate the projected text and images will be run using Unity on the Raspberry Pi, and the LCD screens will act as output monitors. In order for our HUD to be ergonomic, we will also include a fan in the system that will reduce the effects of thermal conduction on the wearer to prevent any discomfort from overheating components. The fan should likely not interfere with other aspects of the helmet's design and is only being used for the proof-of-concept prototype. Acoustic noise from this fan will be negligible compared to the main helmet ventilation flow.

B. Mount

To compile the aforementioned individual components into a particular configuration, we will also need to make a mount such that the system can be easily attached to the astronaut's helmet and adjusted for each user. The system is going to project the image onto the visor such that the image reflects off and is clearly visible to the user. It is important we create an adjustable mount because different users' eyes will be positioned in slightly different locations relative to the helmet. Furthermore, if the image is not directly reflected to where the eyes are positioned, the astronaut will not be able to see the display.

In order to make an adjustable mount, we will use a ball and socket joint which will allow the system to tilt in all directions. To attach the joint to the helmet, we will additively manufacture a compatible adapter on the inside of the helmet. We will then attach the screen and lenses to the mount by screwing them in place. On the other side of the ball and socket joint, we will additively manufacture an attachment for the projection components such as the LCD screens, lenses, and electronics. A preliminary computer aided design image of the concept is shown in Figure 7.

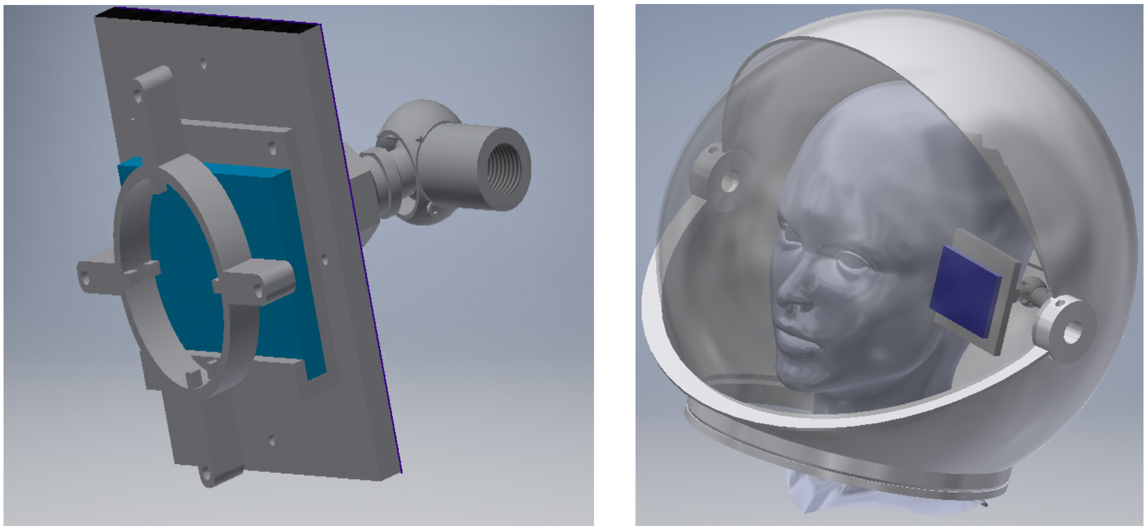


Figure 7. Current Design of Mount (Note: only one mount shown in helmet).

While exploring the design of our HUD projection system, some limitations did arise. For example, the design is constrained to fitting within the helmet while still allowing a comfortable amount of space for the user's head. This limits potential screen sizes we can use and where we can mount the system. Another limitation of this design is that the computer system being used for this HUD is located outside the helmet. As a result, extra wiring must be installed, which introduces new hazards to the suit environment. The computer system itself also presents some limitations as it requires additional components. It should be emphasized that the MX-D suit this helmet is designed for is not pressurized, but uses soft goods tailoring to represent the bulk and restriction of a pressurized suit. This relieves us of the need to make physical and electrical connections to the helmet airtight.

C. Optics

Currently, our planned optical configuration for the heads-up display consists of two TFT (thin film transistor) screens with collimator lenses mounted immediately in front of them projecting onto the visor of the helmet before being reflected back into the wearer's eyes. The collimator lenses will be used to adjust the output angles of the screens so that the resulting image being viewed is not blurry. Due to the curvature of the visor, the screens will project onto opposite eyes; the left screen will project onto the right eye, and vice-versa. We plan on feeding identical images to both screens, though if this makes it hard for the wearer to view the projected image, we may offset one slightly. In preliminary optics testing, we found that adding a double convex lens to the optical path of an off-the-shelf pico-projector is sufficient for the projected image to be read off the acrylic visor at close range. It increases the clarity of the image dramatically compared to the image from the projector alone.

For the purposes of this experiment, we do not anticipate the helmet size changing. However, to account for differences in users, we are including some adjustability in the screen mount in order to ensure the display is focused for different users. In finding the appropriate location for the mount, we used a computerized model of the average human head size and found that the distance from the screen to the display will be about 7.2 in. This was verified experimentally. Therefore, we will use a combination of double convex and plano-convex lenses with a focal length of 7.2 inches. These values and design would need to change in the event the general helmet design is altered for different environments. However, this is outside the scope of this project and will not be discussed or investigated.

We have investigated the effectiveness of several different types of screens at displaying an image projected at short range. We tested anti-glare screens, rear projection screens, and the unadorned acrylic visor for qualities such as transparency (such that written words could be read through it) and visibility, clarity, and brightness of the projected image. Several testers found that the acrylic visor alone exhibited the best qualities for projection inside our helmet, so our optical path will most likely not incorporate extra coatings or screens to boost visibility of the projected image.

There are a few limitations we must consider in this design. First, since the path from the screens to the viewer's eyes is relatively short, we will have to make sure our lenses are perfectly aligned so the projected images are aimed properly. If not, the images may appear out of focus. Second, since different people have differently-spaced eyes, we will need to account for how adjusting the angle of the screens to accommodate them affects the optical path. We are currently relying on the fact that this adjustment should be small, and as mentioned any necessary adjusting will be built into the mounting mechanism itself. Third, we may need to install additional backlighting for our screens to boost their brightness so the projected images do not become washed out by ambient light. While our initial testing indicated that a small backlight should be sufficient, we may need to revise this if we encounter problems with the finished prototype.

D. Future Directions

There are additional accommodations we may explore if feasible within the scope of the project. One of these may include integrating voice commands. This would be fairly simple from a hardware perspective: we would need to add a small microphone to the electronics array within the helmet. The main barrier to including this element would be ensuring that the voice commands would be reliably detected by the microphone within the unusual acoustic environment of the helmet. In addition, we may also incorporate a speaker that will integrate easily with the communications carrier assembly, which is the current audio method used in spacesuit applications. It is difficult to hear voices from outside the helmet when wearing it, so adding a small speaker to our electronics array would significantly improve communication between the people testing out the helmet and other experimenters.

When considering the backlight for our LCD screen, we will initially choose a level of lighting that is appropriate for the environment in which we will do bench and human factors testing. However, in the future we may research the level of ambient lighting that would be expected for astronauts on the Moon or Mars. The ambient brightness in those two scenarios would undoubtedly be different than that of our lab; our projected HUD would then require different levels of backlighting to be visible against the background.

Once we have a working prototype of our helmet-mounted HUD, we intend to repeat the human factors testing we previously described for the HoloLens. This will allow us to compare responses to the two styles of heads-up displays when used to perform the same tasks. We hope that the helmet-mounted display we design will be easier and more ergonomic for EVA-analogue tasks than the HoloLens, but either way our results will be useful for future developments of heads-up displays integrated into spacesuits.

Additionally, though our initial prototype is intended to be a proof of concept just to show a HUD is implementable, if we have time we may begin development on a refined version that would be closer to a flight-ready helmet. Changes we would make for this could include a switch to water cooling for the wearer and the helmet electronics, higher-resolution screens, and an automatically-adjusting backlight, among others. However, these are of relatively low priority at this point since our first goal is making sure the displayed instructions themselves are actually beneficial to the wearer.

V. Conclusion

The exciting recent developments in HUD technologies beg to be incorporated into space suits. The utility is clear: astronauts have limited mobility and access to knowledge in their suits during EVAs, so a hands-free method of delivering written instructions allows them to retain full use of both hands while being able to reference back to instructions and see visuals, not just hear instructions over a headset. This method also mitigates concerns about time delays in communication as humanity begins to explore Mars and beyond. The feasibility of this project is also apparent when one considers that an astronaut helmet already has a visor that can be used as part of an optical pathway, and that advances in pico-projector-related technologies allow for the complete projecting path to be laid out within the confines of the helmet. In our research, we are capitalizing on these advantages as we work towards prototyping and testing heads-up displays inside astronaut helmets.

In our initial testing of the HoloLens HUD, it was found that the HUD did not significantly reduce task completion time. The qualitative results from the testing will inform our design process for a prototype heads-up display that will be built into an astronaut helmet. Our current plans are to use additive manufacturing to build a helmet with an adjustable mount for a projector system, which will then project text and images straight onto the visor. The more precise optical path design details are still ahead of us, as are the incorporation of feedback from HoloLens testing and, most likely, many more iterations of the design. However, we are confident that in the coming phases of our research, we will produce a prototype helmet-mounted HUD that will increase the ease with which astronauts can perform EVA tasks.

VI. Appendix

Two-Factor ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Rows	232355	29	8012.231034	2.518324164	0.001386414	1.6629
Columns	27042	2	13521.00048	*4.24978537*	0.018952452	*3.15593*
Error	184531	58	3181.572551			
Total	443928	89				

Audio Vs. Hololens

	Audio	Hololens
Mean	339.807	333.584
Variance	2696.56	8352.44
Observations	30	30
Pearson Correlation	0.28837	
Hypothesized Mean Difference	0	
df	*29*	
t Stat	0.37386	
P(T<=t) one-tail	0.35561	
t Critical one-tail	1.69913	
P(T<=t) two-tail	0.71123	
t Critical two-tail	2.04523	

Audio Vs. Both

	Audio	Both
Mean	339.807	300.322
Variance	2696.560891	3326.37
Observations	30	30
Pearson Correlation	0.72690555	
Hypothesized Mean Differen	0	
df	29	
t Stat	5.294074831	
P(T<=t) one-tail	5.60128E-06	
t Critical one-tail	1.699127027	
P(T<=t) two-tail	1.12026E-05	
t Critical two-tail	2.045229642	

Hololens Vs. Both

	Hololens	Both
Mean	333.584	300.322
Variance	8352.44	3326.37
Observations	30	30
Pearson Correlation	0.2438	
Hypothesized Mean Difference	0	
df	29	
t Stat	1.90891	
P(T<=t) one-tail	*0.03311*	
t Critical one-tail	1.69913	
P(T<=t) two-tail	0.06621	
t Critical two-tail	2.04523	

Figure 1. ANOVA and T test on times

Test Subject	Audio	Hololens	Both	*Block Average*
	Time (s)	*Time (s)*	*Time (s)*	*Average*
1	359.63	397.60	369.36	375.53
2	325.95	271.47	344.07	313.83
3	379.49	356.59	281.33	339.14
4	390.59	437.10	325.87	384.52
5	250.65	302.47	213.57	255.56
6	328.81	271.43	328.66	309.63
7	366.31	299.61	313.85	326.59
8	323.30	249.39	346.05	306.25
9	280.61	291.14	262.43	278.06
10	412.62	479.26	289.96	393.95
11	369.03	267.08	261.56	299.22
12	341.21	281.57	361.26	328.01
13	481.15	471.74	426.40	459.76
14	315.70	225.60	263.87	268.39
15	404.38	303.86	347.11	351.78
16	413.22	278.76	319.91	337.30
17	350.78	312.68	362.22	341.89
18	413.73	478.22	423.78	438.58
19	342.18	337.57	308.11	329.29
20	316.81	347.05	267.34	310.40
21	283.12	265.43	220.25	256.27
22	324.23	264.74	256.58	281.85
23	277.26	311.5	227.66	272.14
24	324.82	328.35	252.85	302.01
25	336.91	330.82	362.23	343.32
26	301.1	315.6	278.83	298.51
27	307.77	259.31	234.35	267.14
28	319.16	248.35	263.6	277.04
29	272.56	375.23	240.45	296.08
30	359.63	648	256.14	421.26
Average	342.42	333.58	300.32	325.44

Figure 2. Table of times n = 30

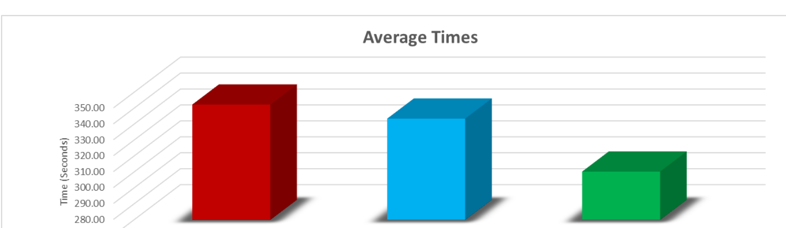


Figure 3. Bar Graph of Average times n = 30

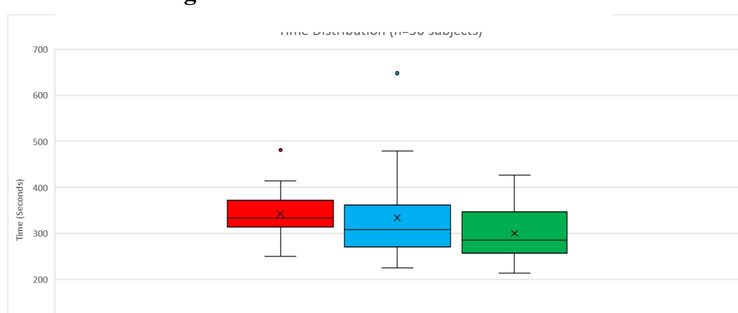


Figure 4. Box Plot of times n = 30

TLX Factor # 1-how mentally demanding was the task?

Test Subjects	Audio	Hololens	Both	Average
1	2	3	2	2.33
2	3	2	3	2.67
3	2	2	2	2.00
4	1	3	1	1.67
5	1	1	1	1.00
6	2	2	2	2.00
7	2	2	2	2.00
8	4	1	1	2.00
9	3	4	3	3.33
10	5	3	1	3.00
11	2	4	3	3.00
12	2	3	2	2.33
13	1	1	1	1.00
14	4	3	2	3.00
15	3	3	4	3.33
16	2	3	1	2.00
17	9	8	5	7.33
18	2	1	1	1.33
19	2	1	2	1.67
20	2	3	2	2.33
21	2	3	2	2.33
22	3	2	3	2.67
23	2	2	1	1.67
24	3	3	3	3.00
25	2	1	4	2.33
26	6	7	5	6.00
27	2	4	1	2.33
28	2	4	4	3.33
29	3	2	1	2.00
30	1	1	1	1.00
Average	2.67	2.73	2.20	2.53

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	151.733	29	5.23218	6.11827957	2.6E-09	1.662900781
Columns	5.06667	2	2.53333	*2.962365591*	0.05957	*3.155931971*
Error	49.6	58	0.85517			
Total	206.4	89				

Hololens Vs.Audio		
	Hololens	Audio
Mean	2.73333	2.66667
Variance	2.68506	2.71264
Observations	30	30
Pearson Correlation	0.71977	
Hypothesized Mean	0	
df	*29*	
t Stat	0.29689	
P(T<=t) one-tail	0.38433	
t Critical one-tail	1.69913	
P(T<=t) two-tail	0.76866	
t Critical two-tail	2.04523	

Audio Vs.Both		
	Audio	Both
Mean	2.666667	2.2
Variance	2.71264	1.544827586
Observations	30	30
Pearson Correlation	0.55588	
Hypothesized Mean Diffe	0	
df	29	
t Stat	*1.81576*	
P(T<=t) one-tail	0.03988	
t Critical one-tail	1.69913	
P(T<=t) two-tail	0.07976	
t Critical two-tail	2.04523	

Hololens Vs. Both		
	Hololens	Both
Mean	2.73333	2.2
Variance	2.68506	1.54483
Observations	30	30
Pearson Correlation	0.63661	
Hypothesized Mean	0	
df	*29*	
t Stat	2.28331	
P(T<=t) one-tail	0.01496	
t Critical one-tail	1.69913	
P(T<=t) two-tail	0.02992	
t Critical two-tail	2.04523	

Figure 5. ANOVA and T test on TLX Mental Demand

TLX Factor # 2-how physically demanding was the task?

Test Subject	Audio	Hololens	Both	Average
1	1	1	1	1.00
2	2	2	2	2.00
3	1	1	1	1.00
4	1	1	1	1.00
5	1	1	1	1.00
6	1	1	1	1.00
7	1	2	3	2.00
8	1	2	2	1.67
9	2	2	2	2.00
10	2	5	2	3.00
11	3	3	4	3.33
12	4	4	4	4.00
13	2	3	1	2.00
14	2	3	4	3.00
15	4	3	4	3.67
16	1	1	1	1.00
17	1	1	2	1.33
18	1	2	1	1.33
19	1	1	1	1.00
20	1	1	1	1.00
21	2	2	2	2.00
22	2	2	3	2.33
23	2	1	2	1.67
24	1	2	2	1.67
25	1	2	2	1.67
26	3	3	3	3.00
27	2	2	1	1.67
28	1	2	2	1.67
29	1	1	1	1.00
30	1	1	1	1.00
Average	1.63	1.93	1.93	1.83

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	67.16666667	29	2.31609	7.6616	3.2E-11	1.6629
Columns	1.8	2	0.9	*2.97719*	0.05877	*3.15593*
Error	17.53333333	58	0.3023			
Total	86.5	89				

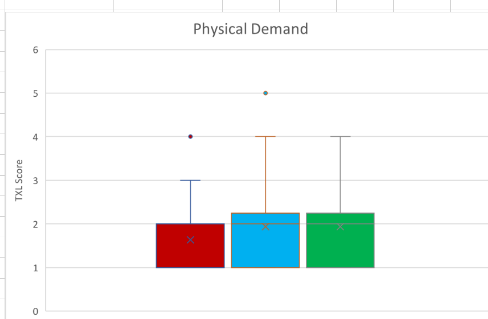


Figure 6. ANOVA and Box Plot on TLX Physical Demand

TLX Factor # 3-how hurried or rushed was the pace of the task?

Test Subject	Audio	Hololens	Both	Average
1	4	7	2	4.33
2	2	3	5	3.33
3	1	1	2	1.33
4	1	2	3	2.00
5	1	1	1	1.00
6	1	1	2	1.33
7	1	1	1	1.00
8	2	2	2	2.00
9	2	3	2	2.33
10	7	1	4	4.00
11	4	5	5	4.67
12	1	1	1	1.00
13	1	1	1	1.00
14	6	4	3	4.33
15	2	2	1	1.67
16	2	3	4	3.00
17	2	2	2	2.00
18	1	1	1	1.00
19	8	8	7	7.67
20	4	6	3	4.33
21	3	4	2	3.00
22	4	4	2	3.33
23	2	4	3	3.00
24	2	2	3	2.33
25	1	1	1	1.00
26	2	4	3	3.00
27	3	2	1	2.00
28	3	5	5	4.33
29	3	3	4	3.33
30	2	2	5	3.00
Average	2.62	2.90	2.62	2.71

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	200.722	29	6.92146	5.55675	1.52756E-08	1.6629
Columns	1.08889	2	0.54444	*0.4371*	0.648019178	*3.15593*
Error	72.2444	58	1.24559			
Total	274.056	89				

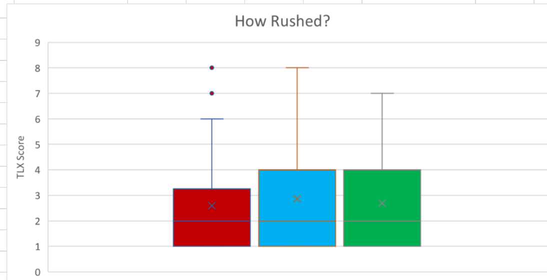


Figure 7. ANOVA and Box Plot on TLX pace

TLX factor # 4-how successful were you in accomplishing what you were asked to do?

Test Subject	Audio	Hololens	Both	Average
1	2	1	2	1.67
2	2	9	2	4.33
3	1	1	1	1.00
4	1	1	3	1.67
5	1	1	1	1.00
6	2	2	2	2.00
7	1	1	3	1.67
8	2	2	2	2.00
9	2	2	2	2.00
10	3	2	1	2.00
11	1	9	1	3.67
12	1	1	1	1.00
13	3	1	1	1.67
14	1	1	1	1.00
15	2	2	3	2.33
16	3	1	1	1.67
17	2	1	10	4.33
18	1	1	1	1.00
19	3	2	1	2.00
20	1	2	1	1.33
21	2	2	1	1.67
22	1	1	1	1.00
23	2	1	1	1.33
24	3	1	1	1.67
25	1	1	8	3.33
26	2	3	2	2.33
27	2	3	2	2.33
28	2	1	1	1.33
29	8	10	10	9.33
30	3	2	2	2.33

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	229.733	29	7.92184	2.53289	0.0013	1.6629
Columns	1.26667	2	0.63333	*0.2025*	0.81726	*3.15593*
Error	181.4	58	3.12759			
Total	412.4	*89*				

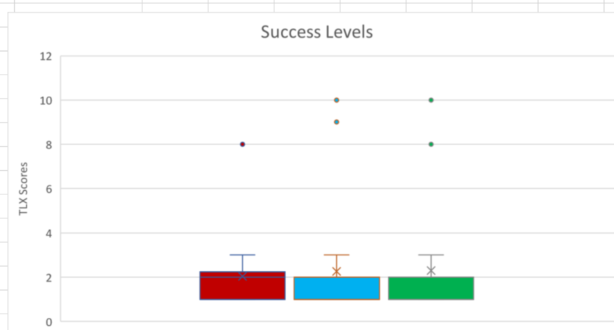


Figure 8. ANOVA and Box and Plot on TLX success

TLX Factor # 5- how hard did you have to work to accomplish your level of performance?

Test Subject	Audio	Hololens	Both	Average
1	1	3	1	1.67
2	2	2	3	2.33
3	2	2	2	2.00
4	1	2	2	1.67
5	1	1	1	1.00
6	2	1	2	1.67
7	1	1	2	1.33
8	3	1	2	2.00
9	3	3	3	3.00
10	6	5	2	4.33
11	3	4	3	3.33
12	2	3	2	2.33
13	3	1	1	1.67
14	5	2	3	3.33
15	4	3	4	3.67
16	2	3	1	2.00
17	8	5	8	7.00
18	2	2	2	2.00
19	2	1	1	1.33
20	3	3	2	2.67
21	2	3	2	2.33
22	3	3	3	3.00
23	2	2	1	1.67
24	2	2	2	2.00
25	2	1	3	2.00
26	5	4	3	4.00
27	1	3	2	2.00
28	4	4	4	4.00
29	3	4	4	3.67
30	2	2	3	2.33
Average	2.73	2.53	2.47	2.58

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	128.622	29	4.435249042	6.09905	2.8E-09	1.6629
Columns	1.15556	2	0.577777778	*0.79452*	0.45665	*3.15593*
Error	42.1778	58	0.727203065			
Total	171.956	*89*				

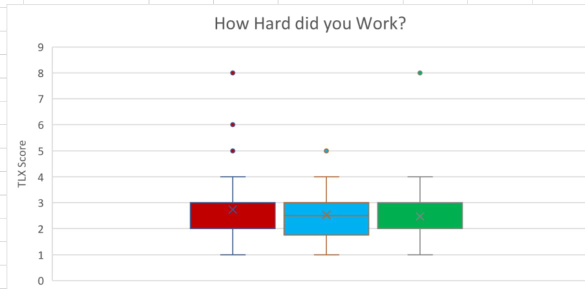


Figure 9. ANOVA and Box Plot on TLX work

TLX Factor # 6- how insecure, irritated, stressed, and/or annoyed were you during the task?

Test Subject	Audio	Hololens	Both	Average
1	3	5	3	3.67
2	1	1	1	1.00
3	2	2	2	2.00
4	1	2	1	1.33
5	2	3	3	2.67
6	2	2	3	2.33
7	2	1	2	1.67
8	3	1	1	1.67
9	1	1	3	1.67
10	6	3	2	3.67
11	2	2	1	1.67
12	1	1	2	1.33
13	5	1	1	2.33
14	4	1	4	3.00
15	1	2	2	1.67
16	2	2	1	1.67
17	3	3	2	2.67
18	2	1	1	1.33
19	4	4	3	3.67
20	2	4	2	2.67
21	2	2	2	2.00
22	5	2	5	4.00
23	2	1	1	1.33
24	1	2	2	1.67
25	1	2	2	1.67
26	2	2	1	1.67
27	1	2	1	1.33
28	1	2	3	2.00
29	1	1	1	1.00
30	2	2	2	2.00
Average	2.23	2.00	2.00	2.08

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	59.7889	29	2.06169	2.23185	0.00468	1.6629
Columns	1.08889	2	0.54444	*0.58938*	0.55796	*3.15593*
Error	53.5778	58	0.92375			
Total	114.456	*89*				

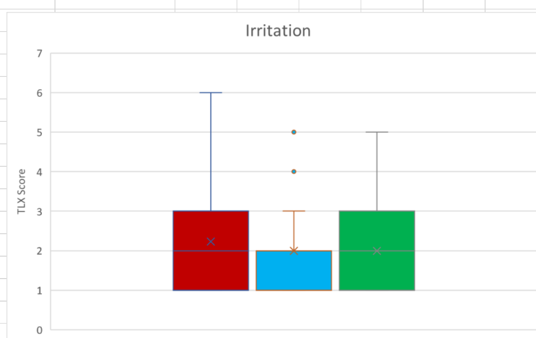


Figure 10. ANOVA and Box Plot on TLX irritation

iii. Cooper Harper

Test Subjects	Audio	Hololens	Both	Average
1	1	2	2	1.67
2	2	3	3	2.67
3	1	4	4	3.00
4	2	2	3	2.33
5	2	2	2	2.00
6	1	1	2	1.33
7	1	2	9	4.00
8	2	1	2	1.67
9	2	2	3	2.33
10	3	2	1	2.00
11	2	3	2	2.33
12	1	1	1	1.00
13	3	2	1	2.00
14	3	1	2	2.00
15	2	2	2	2.00
16	2	1	2	1.67
17	4	4	3	3.67
18	2	2	1	1.67
19	3	2	3	2.67
20	1	3	2	2.00
21	2	3	2	2.33
22	2	2	2	2.00
23	3	2	1	2.00
24	1	2	2	1.67
25	1	2	3	2.00
26	2	2	2	2.00
27	1	3	1	1.67
28	2	3	2	2.33
29	3	2	2	2.33
30	2	3	3	2.67
Average	1.97	2.20	2.33	2.17

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	33.1667	29	1.143678161	0.98612	0.5029	1.6629
Columns	2.06667	2	1.033333333	*0.89098*	0.41579	*3.15593*
Error	67.2667	58	1.159770115			
Total	102.5	89				

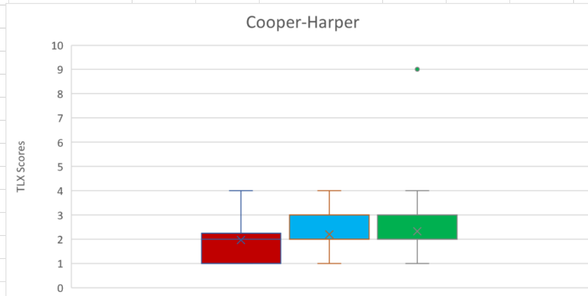


Figure 11. ANOVA and Box Plot on Cooper Harper responses

Acknowledgments

Our team would like to thank Mr. Lemuel Carpenter, for his guidance and support during the early stages of the project. We would also like to acknowledge Dr. Kristan Skendall, Dr. Frank Coale, and the rest of the Gemstone Honors College Staff for their assistance and financial support, and to thank our team's librarian, Lindsay Inge, for her help with written works and data management.

References

[1] Parnell, Brid-Aine, "NASA Won't Have New Spacesuits For Years - Despite $200 Million Investment," Forbes, Apr 28, 2017

[2] Williamson, S. J., and Cummins, H. Z., *Light and color in nature and art*, New York: John Wiley & Sons, 1983.

[3] Williams, L. J., and Abdi, H., "Fishers Least Significant Difference Test," *Encyclopedia of Research Design*, 2010, pp. 840–853.