CHAPTER I

INTRODUCTION

Teleoperation of a mobile robot is a basic operation that requires human–robot interaction. Most current teleoperation methods use hand held controllers, such as joysticks [1, 2, 3, 4]. This thesis presents a Personal Digital Assistant (PDA) based teleoperation interface for a mobile robot. This interface differs from previous PDA teleoperation interfaces [1, 5] in that it uses the PDA as a touch screen. In this interface, three touch-based PDA screens were developed: a Vision-only screen, a Sensor-only screen, and a Vision with Sensory Overlay screen.

There is a desire to provide small, lightweight mobile interaction devices. Currently when robots are deployed for military or search and rescue situations, the common interface is a large, heavy operator control unit similar in size to a toolbox such as those available from Foster-Miller [6] and MESA Associates, Inc. [7]. In addition to heavy operator control units some companies, such as Foster-Miller [6] and Cybernet Systems Corporation [8], have developed wearable operator control units. A focus in providing small, lightweight mobile interaction devices includes the use of PDA’s. Most of the prior work focuses on stylus-based interfaces. Stylus-based interaction is not practical for all users in military or search and rescue domains. A stylus-based interface is difficult to use while walking or running, therefore the focus of this work has been touch-based interaction.

The research problem of this work is to find a good way of presenting the robot’s environment and sensory information to the user while employing a Personal Digital
Assistant (PDA) as the interaction device. The Personal Digital Assistant (PDA) based
teleoperation interface is integrated with a mobile robot in the Intelligent Robotics Lab at
Center for Intelligent Systems in Vanderbilt University. A user evaluation was conducted
to determine which interface screen is the most understandable and facilitates the user’s
decision-making process. Thirty novice users participated in the evaluation in which they
teleoperated a mobile robot using all three interfaces. The perceived workload; the
demographic information; the subjective usability data; the number and location of
touches; the completion times; and the goal achievement accuracy were collected. The
following hypotheses were defined for the user evaluation:

1) The Vision with Sensory Overlay screen and Sensor-only screen induce higher
workload levels for the defined user group than the Vision-only screen.

2) The Vision-only screen is easier to use for the defined user group than the
Vision with Sensory Overlay screen and Sensor-only screen.

3) The Vision-only screen is the preferred teleoperation interface by the defined
user group.

4) The Vision with Sensory Overlay screen is preferred for teleoperation over the
Sensor-only screen by the defined user group.

This thesis presents the motivation and design of the interface and discusses
issues that arise when employing a PDA as a touch screen and the results of the human
factors evaluation. Chapter 2 provides the literature review of PDA based interfaces both
in general and specifically as applied to Human Robot Interaction (HRI) while Chapter 3
provides background information related to this work. Chapter 4 describes the PDA
interface design considerations and provides the screen designs. Chapter 5 explains the
user evaluation experimental design, and Chapter 6 provides the user evaluation analysis and results. Finally, Chapter 7 provides the conclusions and the future work.
CHAPTER II

LITERATURE REVIEW

Personal Digital Assistants (PDAs) are important tools in today’s information technology world. A significant amount of research related to the development of PDA-based interface applications has been conducted. This chapter provides the literature review of the research related to developing general PDA-based interfaces as well as PDA-based interfaces for Human Robot Interaction (HRI).

Standard PDA Interfaces

PDAs generally provide standard interfaces for tasks such as recording addresses, calendar activities, listening music, checking e-mails. In the last few years, many researchers have focused on the development of other capabilities for these devices. Examples of such applications include:

- recognizing characters or signs in specific languages
- collecting daily data
- in educational systems
- for face recognition
- for Internet based mobile surveillance system
- connecting multiple PDAs to a Personal Computer
- using the PDA as an electronic guidebook
- creating a speech – centric multimodal interaction
• evaluating a gesture design tool
• developing user interfaces for different computing platforms

This section focuses on these general, non-robotic interfaces.

A PDA can be used for recognizing characters or signs in specific foreign languages. Luo and Wu [9] designed and developed a PDA interface for recognizing Chinese handwritten characters. Kang and Kim [10] present a system using a PDA to recognize Korean characters. Zhang et al. [11] developed an approach for a PDA-based sign system that presents a sign translator to the user. “A sign can be a displayed structure bearing letters or symbols” [11]. With this system, the PDA can capture an image of a sign. The image is then auto segmented to find Chinese characters and the system then provides the English translation. Figure 1 shows the interface for this system and an example of sign character detection and segmented character images.

Figure 1: (a) Sign Character Detection (b) Segmented Character Images (c) Sign Detection recognition and translation on PDA [11]
The difference between these systems and the PDA-based teleoperation interface developed for this work is that the PDA interface designed for recognizing foreign language requires stylus-based interactions. The work in this thesis focuses on developing a touch screen interface. Therefore, the interface in this work does not require a stylus.

A PDA can also be used for collecting ideas and information while users are not in front of a personal computer (PC). The data can be recorded anywhere that the PDA can be used. Yoshino et al. [12] developed *GMemo* that collects data via a PDA. The design policy was based on the idea of writing down users’ ideas during the day when they were away from a PC. *GMemo* is input equipment for GUNGEN. GUNGEN is a groupware for a new idea generation support system. The concept of GMemo and a supporting flowchart made by GUNGEN using a PDA are shown in Figure 2.

![Figure 2: (a) Concept of GMemo and (b) A supporting flowchart made by GUNGEN using a PDA [12]](image)

This system requires stylus-based interaction, which is different from the work that is presented in this thesis. The data is collected via handwriting input or one-touch stylus input. The system uses handwritten character recognition. The recording location
and the information source can be selected from pop up menus with one-touch stylus input. An example screen of this system is shown in Figure 3.

![Figure 3: An example of the GMemo screen](image)

PDAs also play an important role in educational systems. Breitbart et al. [13] designed Pocket-IMPACT, a web-based educational product used with PDAs. This system presents instructional tools used by students and instructors, such as reviewing a lecture, taking a quiz, and entering or checking grades. Vila et al. [14] developed a teacher interface for a computer supported educational system. Chen et al. [15] created a PDA/GPS development platform for self-guide systems that provides functions for the development of user defined guide systems. They applied this system to native education of a city in Taiwan. This paper defines “native education” as a special type of education, since teaching objectives are different from city to city in Taiwan. Another example of using PDAs in educational system is provided by Avanzato [16]. Twenty-four students in
the Digital Systems course at Pennsylvania State University-Abington used PDAs in the classroom and laboratory. These systems all require stylus-based interactions.

PDAs have also been used for face recognition. Yang et al. [17] developed a PDA-based face recognition system that used a PDA with an attached camera. This system is composed of an interface and a face recognition module. The interface allows the user to select a face and facial features. The face recognition module provides image preprocessing and face recognition algorithms. The system uses the Microsoft embedded development tool on a desktop computer and then downloads to a Pocket PC. The PDA and the camera that are used are shown in Figure 4.

![Figure 4: HP Jornada pocket PC and HP pocket camera [17]](image)

This system requires stylus-based interactions. This is a difference between this work and the work presented in this thesis. The similarity is that both do image processing on a computer and then download the information to a PDA.

The PDA is employed in mobile surveillance services as developed by Li et al. [18]. The goal is to allow users universal access to surveillance services any time from anywhere. The client product is called “PDA Watch”. The PDA Watcher system is shown in Figure 5.
The PDA-based teleoperation interface developed in this work differs from the described PDA Watcher interface in that it uses the PDA as a touch screen. Therefore, the interface does not require a stylus. The PDA Watcher interface requires stylus-based interaction.

Myers et al. [19] have done some interesting research that connects multiple PDAs to a Personal Computer. They developed two main applications: Remote Commander and PebblesDraw. The Remote Commander allows the PDAs to control the PC and multiple co-located users to take turns controlling the PC mouse and the keyboard without leaving their current locations. To move the cursor on the PC, the user moves the stylus or her/his finger across the PDA screen. Therefore, either stylus based interaction or touch based interaction system can be used for moving the cursor in this system. The Pebble’s Remote Commander application is shown in Figure 6.
The *PebblesDraw* application permits multiple individuals to draw at the same time and is a multi-cursor drawing program in which all the users share the same display. The interaction mechanism for this application is stylus-based. The first version of *PebblesDraw* is shown in Figure 7.

![Figure 6: The Palm Pilot running the Pebbles Remote Commander application][19]
Haneef and Ganz [20] have used a PDA as an electronic guidebook. The MANAS system enables users to employ their PDAs as mobile electronic guidebooks. This system provides network access to mobile device users in public places, such as museums. They use PDA as a part of the MANAS system. The MANAS Service Architecture is shown in Figure 8. This system uses stylus-based interaction.
Huang et al. [21] created a speech – centric multimodal interaction framework, called MiPad. MiPad is an application prototype that demonstrates advantages for wireless PDA devices. MiPad integrates continuous speech recognition and spoken language understanding while allowing users to accomplish many common tasks such as e-mail, voice-mail, calendar, contact list, and web browsing. MiPad focuses on using pen and speech-interaction in concert. With MiPad’s Tap and Talk interface, there is no need to use a stylus and PDA’s keyboard. This system uses push-to-talk control. For example, the user may use MiPad to read an e-mail that contains someone’s contact information. The user can tap the Contact field and speak “John Doe’s phone number is 2404567”. This action adds John Doe’s number to the contact list. The MiPad system is shown in Figure 9.
Huang et al. conducted a user study to compare their Tap and Talk interaction with normal pen-based interaction. They focused on task-completion time and user satisfaction. Each participant completed half of the tasks with MiPad interface and half of the tasks with the regular pen-based interface. They found that “simple actions are very suitable for pen-based interaction” and “15 of the 16 participants stated that they preferred using tap and talk interface for creating new appointments and all 16 said that they preferred it for writing longer e-mails.”

The difference between this work and the work presented in this thesis focuses on the MiPad’s use of the combined pen and speech interaction. The PDA based teleoperation interface focuses only on touch-based interaction and does not employ a pen or speech interaction capabilities.

Long et al. [22] designed and implemented a gesture design tool, called \textit{gdt}. Their goal was to improve gesture recognition for stylus based user interfaces. The main window of \textit{gesture design tool (gdt)} is shown in Figure 10. They conducted a user study employing a PDA to evaluate their gesture design tool. They found from the user study
that “users think gestures are powerful, efficient and, convenient”. It was also found that “users want more gestures”. In addition, “users want to define their own gestures, providing a macro-like capability”. Finally, it was found “users are dissatisfied with the recognition accuracy of gestures and they do not want to sacrifice recognition for the sake of more gestures.”

Long et al.’s work is different from the work presented in this thesis in that their work requires stylus-based interaction while the work presented in this thesis requires touch-based interaction.

PDAs have also been used as a platform for designing user interfaces. Eisenstein et al. [23] designed and developed user interfaces for different computing platforms,
ranging from workstations to PDAs and cellular phones. They point out that each computing platform has its own constraints. Some devices are immobile while others are mobile (e.g., PDA). Some systems support extensive graphical capabilities (e.g., a large monitor), while others only provide limited interaction capabilities (e.g., a cellular phone). Some systems are equipped with enhanced input and output devices (e.g., a trackball), while others are constrained with limited input. They indicate that mobile computing increases the probability of environmental change while the user is carrying out a task.

This thesis work was inspired by their work related to different computing platforms for different conditions. The difference between their work and this thesis is that they designed different interfaces in order to accommodate different devices for a task but this work designed different interfaces only for a PDA.

This section focused on general, non-robotic interfaces. A considerable amount of research related to the PDA-based human robot interaction interface applications has been conducted. The next section focuses on PDA-based human robot interaction interfaces.

PDA Based Human Robot Interaction (HRI) Interfaces

Some researchers have focused on the development of PDA-based human robot interaction interface. Examples of such applications include:

- *PdaDriver* interface by Fong [5]
- A multimodal interface for mobile robots by Perzanowski et al.[24]
- *PocketCERO* interface by Huttenrauch and Norman [25]
- Low-bandwidth video and control of a mobile robot on a wireless PDA interface by Ferworn et al. [26]

- PDA based sketch interface by Bailey [27]

This section focuses on these PDA-based human robot interaction interfaces.

Fong [5] designed and implemented the control architecture, collaborative control. His design enables the operator to communicate with and assist the robot. The associated PdaDriver interface permits both collaborative control and vehicle teleoperation.

This interface has four control modes that are shown in Figure 11. The direct mode, image mode, and map mode enable the user to generate robot motion commands and to navigate the robot. The sensor mode allows the user to control sensing and perception. The direct mode allows the user to control the relative position of the robot (pose mode) and the speed of the robot (rate mode). The image mode allows the user to view images from a robot-mounted camera. This mode provides the user with the ability to pan and tilt the camera. The map mode allows the user to view a map in either robot (local) or world (global) coordinates. The robot controller creates the map using a 2D histogram-based occupancy grid and sensor range data. The sensor mode allows the user to control the robot’s camera position and imaging settings as well as sonar array. The user can directly modify the camera gain settings, such as backlight compensation with this interface.
Figure 11: *PdaDriver* Control Modes (a) Direct Mode (b) Image Mode (c) Map Mode (d) Sensor Mode [5]

Fong’s work is different from the work in this thesis in that his interfaces are designed specifically for the collaborative control architecture. This thesis work relies on the user to act as a teleoperator with complete control of the system. The collaborative control can provide benefits over traditional teleoperation. Collaborative control helps humans and robots to better interact and attempts to make human-robot interaction more
like human-human interaction. Moreover it makes teleoperation more flexible, by providing robots greater freedom to act and to function as partners [16]. While it is true collaborative control is intended to assist the operator, the required screens, as designed by Fong, require stylus-based interactions. His screens cannot be easily transformed to accept touch-based interactions due to the limited PDA screen real estate. The PDA-based teleoperation interface developed in this work differs from the Fong’s work in that it uses the PDA as a touch screen and does not require a stylus.

Perzanowski et al. [24] designed and implemented a multimodal interface for mobile robots. The robots understand speech, hand gestures, and PDA-based input. Through the PDA, operators can directly provide a limited set of commands. Figure 12 shows their interface. Operators can command a robot or a team of robots to navigate to a location via a combination of inputs.

Figure 12: A Palm Pilot showing a map of the robot’s environment [24]
The difference between their work and the work in this thesis is related to the interaction capabilities. Their interface has been specially designed to permit multimodal interactions such as gestures and speech interactions in addition to the PDA based interactions. Their research reports indoor laboratory environments that are amenable to such interactions. The focus of the work presented in this thesis, is the development of an interface for military or search and rescue personnel during field operations. These field environments may hinder the ability to use speech and gesture based interactions. Therefore, the goal is to provide the operator with an interface that is available in all potential environments, this fact limits the interaction capabilities to those available directly through the PDA.

Stylus-based interaction is not practical for many users as it is difficult to use a stylus while walking or running. Additionally, a stylus is quite small and can easily be lost or misplaced. Gesture-based interaction is also not practical for all situations as gesture recognition relies on cameras that are sensitive to lighting conditions and field of view limitations. Speech-based interaction is not always practical, as it requires a verbal communication link between the participants; relies on voice recognition technology; and may not be feasible in all situations. Providing voice commands in loud, noisy environments as well as situations when no noise is permitted hampers the use of speech-based interfaces.

Huttenrauch and Norman [25] implemented an interface, called PocketCERO. This interface provides different design prototypes for a service robot that is to be used in home or office environments. They stated that since the service robots are mobile, this mobility must be taken into account for the associated interfaces. Therefore, they
recommended that PDAs could be used as a mobile device as a part of the robot interface components. A prototype robot that helps users with daily transport tasks in an office environment, an animated character (named CERO) on the robot platform, and a graphical user interface were built by Huttenrauch and Norman.

They implemented three different design prototypes for the PocketPC platform. The first design prototype was a map based *Graphical User Interface (GUI)*, the second one was a *one-hand Graphical User Interface*, and the last one was a *standard Graphical User Interface*. All three are shown in Figure 13.

![PocketCERO](image-url)  

(a) Map Based GUI  
(b) One-hand GUI  
(c) Standard Design GUI

Figure 13: *PocketCERO* (a) Map Based GUI (b) One-hand GUI (c) Standard Design GUI [25]

The similarity between their work and the work presented in this thesis is the idea of three different PDA-based interfaces styles. Another similarity is that they also used Microsoft Embedded Visual Basic program. The *Map Based GUI* and *Standard Design GUI* require stylus-based interaction while the *One-hand GUI* requires touch-based interaction. Therefore, their work is different in that the interfaces designed in this thesis
are all touch based, but two of their interfaces are stylus-based and one interface is touch based. Another difference is that their interface does not display sensory data, but this thesis’s interfaces always display sensory data.

Ferworn et al. [26] developed a prototype service delivery model that allows real-time control of a robot (N-CART’s MAX project robot) and allows low-bandwidth video to be generated for and displayed on a PDA. The MAX video server streams the JPEG images through a conversion process on the PC-based wireless video server. The images then resized to 48x48 pixels and converted into monochrome. The wireless video server processes the video to a reduced image size and places it into a compatible format for the PDA. The wireless video server listens for connection request from the PDA client and then retrieves the video stream from the MAX video server. They developed a wireless robot allowing teleoperated control via a web browser communicating over an IP network and used the PDA as the target-controlling device. The video on the PDA is shown in Figure 14.
The difference between their work and the work presented in this thesis is that their work provides the user with only the image information whereas this work presents the user with the potential to view the image, sonar, and laser range-finder information. Another difference is that their work is stylus-based, but the work presented in this thesis is touch-based. The similarity between their work and the work presented here is that the images are processed in a computer and then transferred to PDA.

Bailey [27] developed a PDA based sketch interface for capturing user drawn route maps on a PDA screen. This interface will be applied to an autonomous mobile robot called Guinness. To evaluate his PDA based sketch interface, a user study was conducted. The sketched maps will be translated into commands for the robot to follow. He conducted a user evaluation to determine the usability of the interface and evaluate the ability to translate the sketches into linguistic terms understandable by humans.
The difference between his work and the work presented in this thesis is that Bailey [27] tries to move the robot with commands that are translated from the sketched maps. The similarity between his work and the work presented in this thesis is that both conducted user evaluations of the interface.

This section focused on the development of PDA-based human-robot interfaces. These PDA-based human-robot interfaces compared with the PDA based teleoperation interface presented in this thesis. Any differences and similarities between these works and the work presented in this thesis were provided as part of this section.

This chapter presented the literature review of the research related to developing general PDA-based interfaces as well as PDA-based interfaces for Human Robot Interaction (HRI). The next chapter provides a detailed explanation of the equipment used for this work.
CHAPTER III

BACKGROUND

This thesis employs a Toshiba Personal Digital Assistant (PDA) and an ATRV-Jr mobile robot. This chapter provides a detailed explanation of the equipment used for this work.

Personal Digital Assistant (PDA) Description

A Toshiba E740 Personal Digital Assistant as shown in Figure 15, was employed for the development and testing of the PDA-based teleoperation interface.

Figure 15: Toshiba E740 PDA
This PDA runs the Microsoft PocketPC 2002 operating system and is equipped with 64 MB of RAM, as well as an Intel PXA250 400 MHz XScale Processor. It has a TFT LCD screen with a 240x320 resolution. This screen works well when used inside or outside. This PDA also has dual expansion slots and built-in wireless communication capabilities. The Toshiba E740 is equipped with an integrated Wi-Fi (IEEE 802.11b) modem.

Embedded Visual Basic (eVB) was used to implement the PDA-based interface described in this thesis. eVB is a development tool included with Microsoft’s Embedded Visual Tools 3.0. This tool can be used to create, debug, and deploy Windows CE applications and is based on Visual Basic.

**ATRV-Jr Mobile Robot**

This work employs an All-Terrain Mobile Robot, ATRV-Jr robot developed by iRobot Corporation. Figure 16 (a) shows the original ATRV-Jr robot as produced by iRobot Corporation and Figure 16 (b) shows *Scooter*, Vanderbilt University’s modified robot.
The sensing capabilities of this robot include a forward facing camera with pan-tilt capability, a forward facing laser range finder, and seventeen sonar sensors. The sonar sensors are located around the robot with five sonar facing forward, five facing out from each side, and the last two mounted on the rear of the robot. Figure 17 provides the sonar configuration. The range the sonar sensors is 2 meters. This robot is also equipped with BreezeNet Wireless Ethernet that provides a maximum bandwidth of 10 Mbps.
Figure 17: Sonar and Laser Range-Finder Ranging Sensors on Scooter

An Intel® Pentium III™ processor with 128 MB SDRAM and a custom enclosure with specially shock mounted hard-drives are integrated into the robot. The system specifications include: an ATX motherboard; 18 GB IDE hard drive; an intelligent high speed multi-port serial card (8 ports); Linux operating system; and Mobility Robot Integration Software.

This robot has been equipped with a SICK Laser Measurement System. This system provides a range of 150 m with coverage of 180 degrees. The sensor has an angle resolution of 0.5% with +/- 50 mm. distance measurement resolution. The range of the laser sensor is 3 meters.
The robot’s imaging system includes a 2 PCI Frame Grabber for the 2XC999 Color CCD cameras with 6 mm lenses.
CHAPTER IV

PDA-BASED TELEOPERATION INTERFACE

This chapter provides detailed information regarding the general PDA-based teleoperation interface and the individual interface screen designs [28].

General Interface Design Considerations

Many issues arise when designing a PDA-based interface. Since the PDA screen size is limited, important information regarding the robot during specific conditions cannot be simultaneously displayed. Interfaces designed for desktop or laptop computers are able to simultaneously display more information and data. The interaction capabilities provided with standard computers include mouse, and keyboard interactions, and potentially speech and gestures. A PDA-based interface provides only limited display real estate and interaction capabilities. Therefore, it is suggested that a touch-based PDA interface composed of different screens for different conditions will facilitate the screen real estate and interaction constraints. For this reason, the PDA-based teleoperation interface is composed of three different screens for different conditions.

The first screen, a Vision-only screen, may be preferred when the user cannot directly view the robot and the associated environment. The second screen, the Sensor-only screen, may be preferred when the user is teleoperating the robot in an area that contains many objects and obstacles. The third screen, the Vision with Sensory Overlay screen, may be preferred when the user needs to simultaneously view all forward facing sensory data and camera images in a single screen.
If computer-based interfaces are compared with PDA based interfaces, some issues arise other than screen size and interaction capabilities. The PDA interface is a small, portable device that is available everywhere, but computer-based interfaces are not as convenient. Even laptop computers are cumbersome to travel with and are not always appropriate in some environments. On the other hand, PDAs have limited software and computing resources, therefore programming a PDA can be more difficult than programming computer-based counterparts.

Issues arise when designing a touch-based interface for a PDA. Touch-based interfaces require buttons and interaction capabilities that are large enough to accommodate human fingers. In certain environments, users in military or search and rescue situations may wear bulky gloves, and interaction capabilities must accommodate the additional size. Therefore, the buttons occupy much of the PDA screen area. In order to maximize the information presentation and accommodate large buttons, transparent buttons are used to provide the ability to see through the buttons for viewing the underlying information. If standard buttons were used, they would block the view and would hide the area behind the buttons.

**General Interface Interaction**

The standard interaction provides five large, transparent buttons, four for driving and one emergency stop button. The four drive buttons are laid out to represent the corresponding direction the robot would travel if the user was standing behind the robot or standing next to the robot while facing the same direction as the robot. The drive button functionality remains constant, even if the user is facing the robot. It is possible
that this fixed functionality may confuse the user, but our focus is military or search and rescue operations. Therefore, it is anticipated that users would receive extensive training and would be able to adapt to this limitation.

As can be seen from Figure 18, the four drive buttons are aligned along the outer edges of the screen. This was done to ensure that the central screen real estate is fully visible. A benefit of positioning the buttons along the edge is that the tactile feedback can be employed with minimal or no visual requirements when providing commands. The position of the emergency stop button was specifically chosen to provide maximal tactile feedback while ensuring the button’s selection could not be recorded as any other command. The corner positioning provides easy access to this button without requiring visual interaction. This particular position also minimizes the required reach distance. A complication appeared when this button was placed in the other three screen corners. The user could accidentally activate system functions rather than issue the emergency stop command.

![Figure 18: The four drive buttons and the emergency stop button](image)

Figure 18: The four drive buttons and the emergency stop button
The robot can be commanded to drive forward, backwards, turn left, turn right, or a combination of forward motion and turning or backward motion and turning. The user selects button 1, in Figure 18, to command the robot to move forward. Button 2 provides the reverse command. Button 3 commands the robot to turn left and button 4 commands the robot to turn right.

If the user commands the robot to move forward (button 1) but over shoots the desired location, simply selecting reverse (button 2) commands the robot to stop moving forward and start moving backwards. A similar behavior exists if the robot is commanded to turn left (button 3) and then the user selects turn right (button 4). Of course, the user could select the emergency stop button before commanding the above mentioned changes but this action would require two selections to command the same action. First the selection of the emergency stop button followed by the next directional selection. The current implementation is intended to reduce the number of required interactions.

If the user commands the robot forward (button 1) but then needs to turn while still moving forward, all that is required is the selection of the appropriate direction (either button 3 or 4). The robot then completes the turn while moving forward. During this behavior, the user may reverse the drive direction (button 2) or change the turn direction (buttons 3 or 4) without halting the robot’s progress. Again, the user could issue an emergency stop command and then select the drive and turn directions. The result would be a total of three-interface button selections rather than one.

Interaction Type

The PDA-based teleoperation interface is designed to be rotated so that the PDA can be strapped to the arm like a wristwatch. In this case, the user does not hold the PDA,
which frees up the user’s hands. The interaction interface and PDA are rotated 90 degrees counter-clockwise so that the PDA can be attached to the arm, as presented in Figure 19.

![Figure 19: PDA alignment.](image)

**Screen Transitions**

The PDA-based teleoperation interface provides three screens. The ability to easily and quickly transition between the screens is important. Since the interface is intended to be strictly touch-based, menus nor the existing graphical button functionality cannot be used for this transition, as that would require fine-grained stylus interactions. One motivation was to develop the interface to require only gross interaction capabilities. Therefore, the PDA’s program buttons have been reprogrammed to provide the screen transition capability.

Figure 20 provides a guide for the ensuing discussion. The first button, labeled 1 in the figure, displays the *Vision-only* screen when selected. The second button, labeled 2 in the figure, when selected displays the *Sensor-only* screen while the third button, labeled 3 in the figure, displays the *Vision with Sensory Overlay* screen. The button labeled 4 provides the ability to close for the displayed interface screen.
The interface designs are based upon providing rapid but meaningful information from a robot. Three screens were developed to employ visual, sonar, and laser range-finder data. An important criterion during the design phase was the presentation of information rather than raw data. It is not uncommon to see raw sonar data simply presented as protruding from a graphical representation of a robot. Such displays have been shown to be difficult to understand by non-engineering users [29]. Therefore, the focus was to design displays that would provide insight into the robot’s view of the world while minimizing the associated data processing requirements.

The display designs are not without disadvantages. Each screen provides limited sensing capability. Only the Sonar and Laser screen provides the user with sensory information all around the robot. The remaining two screens only provide sensory information from the front of the robot. This limitation leaves the robot susceptible to
undetected attacks in military situations. It also hinders the user’s ability to back the robot up long distances. Such capabilities may be required when the robot is in a confined space and cannot turn around. This is an issue that will be addressed by future interface design refinements. The following sections provide detailed descriptions of each interface.

Vision-only Screen

Purpose

The purpose of the Vision-only screen is to provide a screen containing only the camera image that allows the user to observe the remote environment.

Screen Description

The Vision-only screen is intended to provide the user with the view obtained by the robot’s forward facing camera and standard interaction capabilities, as shown in Figure 21. The camera provides no pan or tilt functionality for this interface screen. Most individuals find themselves comfortable working with real-time image based systems because of the similarities to their own visual system. As well, images typically provide a large amount of information that is not easily obtained with other sensory modalities.

As can be seen in Figure 21, the Vision-only screen provides the user with an image that encompasses the entire screen. The transparent buttons permit the underlying image to be viewed.
It is intended that a user should be able to command the robot even when the robot is not within the human’s sensory field. In other words, the robot is located at a remote location. In this case, the user must rely upon the robot’s sensing capabilities in order to understand the remote environment. An image display of the remote robot’s environment is a logical choice.

Implementation Details

The camera image as obtained from the camera is 640 x 480 pixels while the PDA screen size is 240 x 320 pixels. Therefore, the image must be resized to the required PDA display size. Since the PDA has limited processor power, this resizing occurs on the robot’s processor instead of PDA’s processor. The resized image is then sent via wireless communications to the PDA for display. A benefit of this design is that the amount of information transferred between the robot and the PDA is reduced since the resized image is significantly smaller than the original image.
Advantages and Disadvantages

The advantage of this screen is the camera image. Humans are very familiar with images because of the similarities to their own visual systems. Additionally, images are able to provide a large amount of information. Therefore, it is natural for humans to view the images and interpret the state of the environment in front of the robot.

The disadvantage of this screen is the forward facing camera image. The user is unable to obtain environmental information pertinent to the sides or back of the robot therefore limiting user’s ability to determine the state of the world in that area.

Another disadvantage may occur in monochromatic images, as found in building collapses, such images make it difficult to distinguish obstacles and environmental details. The image shown in Figure 22 was taken by a robot’s camera during a search and rescue response at the World Trade Center [30]. As can be seen from this image, it is difficult to distinguish objects and environmental details.

Figure 22: An image taken by a robot at the World Trade Center [30]

An additional disadvantage is related to the image provided. The position of the camera and the corresponding camera parameters limit the field of view such that an
unviewable area exists between the provided image and the front of the robot. Therefore, it is possible that obstacles located in this unviewable area may be undetected by the user.

Sensor-only Screen

*Purpose*

The purpose of this screen is to provide an interface screen that allows the user to see all detected objects surrounding the robot by employing all available ultrasonic and laser range finder information.

*Screen Description*

The *Sensor-only* screen allows for the detection of objects and obstacles around the entire robot during teleoperation, as shown in Figure 23. This screen should permit the user to command the robot even when the robot is located in an area that contains many objects and obstacles.

Sonar and laser range finder information is not necessarily intuitive to non-engineering users. This may hinder the user’s ability to command the robot. An attempt has been made to provide a more meaningful display that provides some perspective to the information detected by the sensors. As can be seen in Figure 23, the sonar data is represented as rectangles that are displayed only for valid sonar readings. The laser information is displayed as outlines to provide an indication of the information detected by the laser.
This screen is best used when the user is able to directly view the robot and its environment. It is not very useful when the robot is located at a remote environment since it does not provide the camera image, thus making it difficult to fully understand the remote environment.

**Implementation Details**

The raw sonar data and laser range finder data are gathered and processed for display. The raw data from the sonar are processed to determine which sonar has valid readings and rectangles are drawn to represent the detected objects. The raw data from the laser-range finder is processed to connect all 180-laser data points to each other. If the reading is not the max distance, the readings are drawn as connected lines in order to represent the detected objects. The processing is done on the PDA’s processor. The raw sonar and laser range finder data points are sent from the robot to the PDA via the wireless connection.
Advantages and Disadvantages

An advantage of the Sensor-only screen is that it provides sensory information on all sides of the robot. Therefore, the user should be able to obtain environmental information pertinent to the sides or back of the robot. This screen is the only screen that provides environmental information surrounding the robot. The remaining screens provide feedback only for the front of the robot.

An additional advantage occurs in cluttered environments since the user is able to detect any obstacles that are close to the robot in this screen. The other two screens have a dead zone between the edge of the camera’s field of view and the robot, this can make the other screens more difficult to use in cluttered environments.

The disadvantage of this screen is that the user may not understand the presentation of the information since it does not combine the sensory information with the camera image information. As previously mentioned, most individuals find themselves comfortable working with real-time image based systems because of the similarities to their own visual system.

Vision with Sensory Overlay Screen

Purpose

The purpose of this screen is to provide an interface screen that allows the user to see the sonar and laser data from the front of the robot combined with the camera image.
Screen Description

The Vision with Sensory Overlay screen provides the user with the sonar and laser range finder data overlaid on top of the camera image and standard interaction capabilities, as shown in Figure 24. The intention is to provide an image that is enhanced by the representation of what the robot detects in the scene via the sonar and laser range finder. Therefore, the sensory data and camera image information is presented in concert and the user is able to simultaneously view all forward facing sensory data.

The location of the large rectangle in Figure 24 represents the object detected by the sonar sensor and the location of small rectangles represents the objects detected by the laser range-finder sensor.

![Figure 24: Vision with Sensory Overlay Screen](image)

Implementation Details

The raw sonar, laser range finder, and camera image data are gathered for processing. As with the Vision-only screen, processing is required to resize the image to the proper PDA display size. This resizing processing occurs in robot’s processor since
the PDA has limited processing power. The raw sonar and laser range finder data are represented similarly to the sensor-based interface with large rectangles representing the sonar data and small rectangles representing the laser range finder data. The processing of the raw data occurs on PDA’s processor. This information is then overlaid on top of the camera image. In this design, the real world object coordinates are mapped to the image plane pixel coordinates using the camera parameters, and camera must be calibrated for measuring the camera parameters. The MATLAB Camera Calibration Toolbox was used to calibrate the camera so that the sonar and the laser range finder data can be overlaid on top of the camera image. The relative positions of the sonar, the laser range finder, and the camera are included in the calculations.

Advantages and Disadvantages

The advantage of this screen is that it provides all forward facing sensory and vision information. This should allow the user to quickly determine what the robot detects via the on-board sensing capabilities. Future versions of the system may include automatic obstacle avoidance in which case, this screen should facilitate the user’s ability to quickly determine if the robot’s sensors detect an obstacle and anticipate that the robot will avoid the obstacle.

An additional advantage of combining an image with sensory feedback is the ability to identify obstacles in monochromatic images as found in building collapses. This identification is possible due to the use of the sonar and laser range finder. As previously mentioned, it is difficult to distinguish objects and environmental details in monochromatic images as shown in Figure 22.
There are three primary disadvantages associated with this screen. First, this screen may be confusing if the user does not fully understand the sensory information presentation. This is due to the overlaying of the sensory data on top of the image. Users are comfortable with interpreting images, but the addition of the sensory information may not be clear.

The second disadvantage is that this screen does not provide environmental information surrounding the robot; it only provides feedback from the front of the robot. As previously mentioned, this can expose the robot to attack and make it difficult to back the robot up for long distances.

Finally, there is a possibility that the sonar and laser range finder can detect obstacles that are not within the camera’s field of view. In particular, due to the position of the camera on the robot, the sensors may detect obstacles in front of the robot that exist between the robot and the viewable camera image. This screen hampers the presentation of such information. If this information is simply ignored, then the user may encounter problems.

This chapter presented the details related to the general PDA-based teleoperation interface and the designs of three interface screens. A user evaluation was conducted. The next chapter provides a detailed explanation of user evaluation experimental design.
CHAPTER V

EXPERIMENTAL DESIGN

User Evaluation Design

As a part of this thesis, a user evaluation was conducted in August 2003. One purpose of the evaluation was to determine which screen is most understandable and facilitates the user’s decision-making process. The second purpose was to investigate the usability of each interface screen.

The user evaluation design was within-subjects as all the participants used each of the three interface screens tested. The dependent variables for this user evaluation were preference ratings, task completion times, number of errors, and ability to reach the goal locations. The independent variable was interface design or features of the interface designs. The independent variable is the only manipulation varying between the treatment conditions. [31]

Thirty novice users participated in the evaluation. No participants had prior experience working with mobile robots but all had experience working with PDAs. The participants completed two trials for each of the three interface screens.

Apparatus

The PDA was used in a location from which the participants were unable to directly view the robot or the robot’s environment. The exception was the second trial with the Sensor task. Figure 25 shows the environmental setup for the evaluation. The diamond shapes represent the start points. The three-color landmarks represent the goal
points. The destination point was the three-color landmark in the evaluation. The light gray cylinders, light gray cylinders, and the pink box represent the obstacles. The light gray cylinders represent the trashcans and the dark gray cylinder represents the storage basket.

Figure 25: The evaluation environment

At the beginning of the user evaluation, the participants signed a consent form agreeing to participate in the user evaluation. The consent form is provided in Appendix A. The participants then completed the pre-experimental questionnaire, provided in Appendix B. After that the participants completed a thirty-minute training session that taught them how to drive the mobile robot using PDA. They were provided information regarding the capabilities of the mobile robot and how to use the interface. After the training session, the participants completed their first task followed by a post-task questionnaire. The post-task questionnaire focused on the task that they had just finished. A copy of the post-task questionnaire is available in Appendix C. The participants then
completed their second task followed by the post-task questionnaire and their third task followed by the post-task questionnaire. After finishing all three tasks and post-task questionnaires, the participants completed the post-trial questionnaire in which they compared the three interface screens. After completing the post-trial questionnaire, the participants repeated the tasks and post-task questionnaires. Once they completed all three tasks, they again completed the post-trial questionnaire. The total time required for the evaluation per participant was approximately two hours. The presentation of the tasks was randomized with a total of five participants completing each combination of task presentations.

Additional data collection including the time required to complete each task, the number of precautions and errors participants made, the location and number of screen touches, and the participants’ ability to obtain the final goal position was collected during the evaluation. The completion time of each task and the location and number of screen touches were collected automatically within the program. The number of precautions/errors and the accuracy to goal achievement were collected by the human out in the hall with the robot.

Task Descriptions

All participants completed each of the three tasks. They were instructed to avoid hitting all obstacles as well as door jams and walls. They were also instructed to avoid obstacles along their path to the goal point. Finally they were asked to try to complete the task in less than five minutes. The participants completed two trials with each interface screen.
The evaluation included three different tasks, one for each interface screen. One task required participants to use the *Vision-only* screen. This task will be referred as the *Vision task*. This task required the participants to move the robot forward until it reached to the corner, and then they turned the robot to the left and moved it forward to a position as close to the destination point as possible. The destination point was the three-color landmark near the “Exit” in Figure 25.

The *Sensor task* required participants to use the *Sensor-only* screen to teleoperate the mobile robot. Again, they moved the robot forward from the start point, as indicated in Figure 25, until the robot reached the corner. At the corner, they turned the robot left and moved it forward until the person in the hallway told them to stop. Since the participants could not view the remote environment with the *Sensor-only* screen they were unable to see the destination point, therefore the person in the hallway told the participants when to stop. The verbal verification was used to indicate arrival at the destination point. During the first trial, participants were located in a remote location separate from the environment and the robot. The only information provided was via the display of the ultrasonic and laser range finder sensors. During the second trial, the participants were permitted to stand in the hallway where the robot was to be teleoperated, thus permitting the participants to directly view the environment.

The *Vision with Sensory task* required participants to use *Vision with Sensory Overlay* screen. The participants again moved the robot forward from the start point until the robot reached the corner. At the corner the participants then turned the robot to the right and moved forward to obtain a position as close as possible to the destination point.
**Questionnaires**

The questionnaires were used to gather subjective data. The pre-experimental questionnaire collected demographic information. The post-task questionnaire gathered NASA Task Load Index (TLX) [32] scale ratings and subjective usability data. The post-trial questionnaire gathered the NASA TLX paired comparisons and further subjective usability data. The post-task questionnaire was administered after the completion of each task. The post-trial questionnaire was administered after the completion of the first trial of the three tasks and after the second trial of the three tasks.

One objective of the evaluation was to measure the participants’ perceived mental workload while completing the tasks. The NASA TLX [32] was chosen as the tool to collect the subjective ratings. NASA TLX requires that participants complete two steps. Participants rate their perceived mental demand, physical demand, temporal demand, own frustration, own effort, and own performance along a scale from 1 to 100. Participants also must select the item from all paired comparisons of the factors that most contributed to their perceived mental workload. Perceived mental workload is determined by a simple calculation. The rating of each factor is multiplied with the weight of that factor. The weight factor is the total number of selections for that factor from all paired comparisons of the factors. The product of all factors are summed and then divided by the total weights which is always fifteen since there are possible combinations of the six factors. The result of the division provides the perceived mental workload.

The pre-experimental questionnaires gathered information regarding the participants’ level of experience using computers, PDAs, and robots. The pre-experimental questionnaire and the associated raw data can be found in the Appendix B.
The post-task questionnaire specifically gathered information regarding participants’ experience while performing the individual tasks with the mobile robot. The post-task questionnaire was prepared with MATLAB; therefore the participants completed the computerized post-task questionnaire on a computer. The participants provided values for each NASA TLX factor on the provided scales. The questionnaire also included four Likert scale questions that used a 5-point scale. The questions concerned the perceived usability of the interface, the ability to interpret and understand the presented information, their level of control over the robot during the task and their ability to correct their errors during the task. The post-task questionnaire and the associated raw data can be found in the Appendix C.

The post-trial questionnaire gathered information regarding the NASA TLX paired comparisons and participants’ rankings of interface screens. The post-trial questionnaire was prepared in MATLAB; therefore the participants completed a computerized version of the questionnaire. There are a total of fifteen paired comparisons for the NASA TLX tool that correspond to all combinations of the six factors. The participants chose the factor that they thought contributed most to their total perceived workload during the three trials. They also answered five ranking questions. These questions asked the participants to rank the three interface screens according to the ease of use, ability to understand the data from the robot, ability to understand what the robot was doing, ability to understand the environment in which the robot was working, and according to the general overall rating. Participants ranked each interface screen as: 1 – lowest rating, 2 – middle rating or 3 – highest ranking. Each value could be used only
once for each question. The post-trial questionnaire and the associated raw data are provided in the Appendix D.

Participants

Demographic Information

Thirty volunteers participated in this evaluation. The demographics of all the participants are summarized in Table 1, complete details are provided in Appendix B. There were eight females and twenty-two males ranging in age from 18-40. Twenty-three participants were in the 18-25 age range, four were in the 25-30 age range, and three were in the 30-40 age range. Twenty-seven participants were right-handed while only three were left-handed.

Table 1: Participant Demographics

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Gender</th>
<th>Handedness</th>
<th>Level of Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-25: 23</td>
<td>Female: 8</td>
<td>Left: 3</td>
<td>High School: 8</td>
</tr>
<tr>
<td>25-30: 4</td>
<td>Male: 22</td>
<td>Right: 27</td>
<td>Bachelors: 17</td>
</tr>
<tr>
<td>30-40: 3</td>
<td></td>
<td></td>
<td>Master’s: 5</td>
</tr>
</tbody>
</table>

According to level of education, seventeen participants had completed a Bachelors degree; five had completed a Master’s degree and eight had only a high school education, seven of them were college students, and one was trying to complete his high school education. One participant was Computer Systems Analyst, one was an Electrical Engineer, and one was a musician. Twenty-seven participants were students. Seven of the
student participants were working on a Bachelors degree; one of the students was working on a high school degree; fourteen were working on a Master of Science degree, and five were working on PhD degrees. Thirteen of the student participants were engineering majors, nine participants were science majors, two were arts and science majors, and two were management majors.

The first pre-experimental question attempted to estimate the participants’ level of daily computer usage. Twenty-five participants indicated that they (a value of five) use computers daily. Three participants rated this question as 4 while only two participants rated their usage as being close to never (a value of two). The average response was 4.70 with the standard deviation of 0.79.

The second question attempted to estimate how frequently participants play computer games. Six participants indicated that they always (a value of five) play computer games. Three participants rated this question as 4, and ten participants rated this question as 3 while five participants rated their level of computer game playing as close to never (a value of two) and six rated this response as never (a value of one). The average response was 2.93 with the standard deviation of 1.39.

The third question attempted to estimate the participants’ level of computer expertise. Thirteen participants indicated that they are computer experts (a value of five). Eleven participants rated this question as 4, and five participants rated this question as 3, while only one participant indicating that he had no computer expertise (a value of one). The average response was 4.17 with the standard deviation of 0.95.

The fourth question attempted to estimate the participants’ level of Personal Digital Assistants (PDA’s) usage. Five participants indicated that they always (a value of
five) use PDAs. Two participants rated this question as 4, nine participants rated this question as 3, while fourteen participants rated their PDA usage as being close to never (a value of two). None of the participants rated this question as 1. The average response was 2.93 with the standard deviation of 1.11.

The fifth question attempted to estimate the participants’ level of experience working with robots. All participants rated their experience level as never (a value of one). The average response was 1.00 with the standard deviation of 0.00.

Hypotheses

Four hypotheses were defined. They are:

- The Vision with Sensory Overlay and Sensor-only interface screens induce higher workload levels for the defined user group than the Vision-only interface screen.
- The Vision-only interface screen is easier to use for the defined user group than the Vision with Sensory Overlay interface screen and Sensor-only interface screen.
- The Vision-only interface screen is the preferred teleoperation interface by the defined user group.
- The Vision with Sensory Overlay interface screen is preferred for teleoperation over the Sensor-only interface screen by the defined user group.

This chapter has presented the details related to the user evaluation design. It provided the pre-experimental questionnaire results, hypotheses, and the evaluation apparatus, which included the task descriptions, questionnaires and participants’ demographics. The next chapter provides a detailed explanation of the results of the user evaluation.
CHAPTER VI

EXPERIMENTAL RESULTS

As previously mentioned a user evaluation was conducted in August 2003 as a part of this thesis. This chapter provides a detailed explanation of the user evaluation results. The first section provides information regarding the analysis of the perceived workload. Then, the second section provides information regarding the analysis of the Likert scale questions that the participants were asked in the Post-Task questionnaire. The third section provides information regarding the analysis of the ranking questions from the Post-Trial questionnaire. The fourth section provides the analysis of the task completion times while the fifth section reports the results related to the number of errors committed. The sixth section provides the location and number of screen touches. Finally, section seven provides the results related to the participant’s accuracy when completing the tasks.

Workload Analysis

This section provides information regarding the analysis of the perceived workload data. The post-task questionnaire gathered NASA Task Load Index (TLX) [32] scale ratings while the post-trial questionnaire gathered the NASA TLX paired comparisons. The post-task questionnaire was administered after the completion of each task. The post-trial questionnaire was administered after the completion of the first trial of all three tasks and after the second trial of all three tasks.
Thirty participants completed the user evaluation; therefore there are 30 data points per task per trial resulting in a total of 60 observations per task. A two-way Analysis of Variance (ANOVA) with repeated measures, within subjects analysis was conducted on this data. Results from this analysis are considered significant if $p \leq 0.05$. If this analysis was significant, the data was further analyzed using a paired samples t-test. Results of this analysis are considered significant if $p \leq 0.05$.

Overall Analysis

The perceived workload descriptive statistics were calculated for all screens and are provided in Table 2. The table indicates that the mean perceived workload for the Vision-only screen increased slightly by 1.36 between the first and second trials. The mean perceived workload for Vision with Sensory Overlay screen also increased by 1.23 between trials. The mean perceived workload for Sensor-only screen decreased dramatically between trials. This decrease was the result of the participant’s environmental condition changing between the trials as described in Chapter 5 in page 47.

Table 2: The descriptive statistics for perceived workload for all screens

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>Mean 43.27</td>
<td>Mean 52.47</td>
<td>Mean 54.55</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 15.25</td>
<td>Standard Deviation 18.18</td>
<td>Standard Deviation 16.70</td>
</tr>
<tr>
<td>Trial 2</td>
<td>Mean 44.63</td>
<td>Mean 37.31</td>
<td>Mean 55.78</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 18.26</td>
<td>Standard Deviation 15.61</td>
<td>Standard Deviation 16.39</td>
</tr>
</tbody>
</table>
The perceived mental workload means for all screens are shown in Figure 26. As seen from the figure, the mental workload mean for *Vision-only* screen increased slightly from Trial one to Trial two, the mean for *Sensor-only* screen decreased dramatically from Trial one to Trial two, and finally the mean for *Vision with Sensory Overlay* screen increased slightly from Trial one to Trial two.

![Figure 26: Perceived Mental Workload Means](image)

The repeated measure ANOVA was conducted on the perceived mental workload data to identify significant results between all three tasks across trials, screens, and screens by trials. The result indicate that the difference between Trial one and Trial two across the three tasks is statistically significant, $F(1, 29) = 10.52$, $p < 0.001$. Given the results of this analysis, further analysis was conducted to compare individual task pairs.
Vision vs. Sensor Task Analysis

The repeated measure ANOVA was conducted on the perceived mental workload data to determine if a significant relationship existed across trials and tasks for the Vision and Sensor tasks. The results indicate that this relationship is statistically significant, $F(1, 29) = 13.736, p = 0.001$. Since this result was significant, a t-test was conducted on the individual tasks within a trial. The t-test results for the Vision versus the Sensor tasks for Trial one was found to be statistically significant, $t(29) = -2.5, p = 0.019$. This result indicates that the difference between the workload values during Trial one for these two tasks was significant. In this case, the workload value for the Vision task was significantly less than that of the Sensor task. Similar analysis for the same tasks during Trial two was also statistically significant, $t(29) = 2.477, p = 0.019$. In this case, the workload value for the Sensor task was significantly less than that of the Vision task.

The ANOVA results comparing the Vision task directly with the Sensor task across both trials was insignificant, $F(1, 29) = 0.142, p = 0.709$. The mean perceived workload for the Vision task was 43.95 (std. dev. = 16.76). This mean is less than the perceived workload mean for the Sensor task (mean = 44.89, std. Dev. = 16.89) but there is little difference between the two means. Therefore, when looking at the data across both trials by task, no difference is detected. This is related to the fact that the second trial of the Sensor task permitted the participants to directly view the environment and the robot, thus dramatically reducing their perceived workload.

An analysis of the Vision and Sensor tasks in Trial one versus Trial two was found to be statistically significant, $F(1,29) = 8.231, p = 0.008$. This result indicates that the perceived workload values were statistically different across the trials. The mean
perceived workload for Trial one was 47.87 (std. dev. = 16.72). This mean is higher than the mean for Trial two (mean = 40.97, std. dev. = 16.94). The average workload in aggregate for both tasks during the second trial was lower. The individual task results indicate based on Table 2; that the value increased slightly for the Vision task during the second trial. It was anticipated that the perceived workload values between trials would decrease as the participants gained more experience with the system. The resulting reduction is not related to this factor. Rather the reduction is primarily due to the fact that Trial two of the Sensor task permitted the participants to directly view the robot. During the first trial participants were located in a remote location separate from the environment and the robot, thus making the task much more difficult. This difference in the task executions between trials created a large decrease in the perceived workload values for the Sensor task, which resulted in a statistically significant result across trials.

This subsection shows that the difference between the workload values across Vision task and Sensor task during Trial one and Trial two was statistically significant. This result shows that the participants were able to complete the Vision task with lower workload than the Sensor task in Trial one, but that this result was reversed when participants could directly view the environment and robot during the Trial two Sensor task. What is not clear from this analysis is if the relationship between these two tasks would remain if participants had been permitted to view the robot and environment during Trial two while completing the Vision task. This variation of the task would be expected to provide lower overall workload values. What is unknown is if these results would show that the Vision task required less workload than the Sensor task.
Sensor vs. Vision with Sensory Task Analysis

The repeated measure ANOVA was conducted on the perceived mental workload data to determine if a significant relationship existed across trials and tasks for the Sensor and Vision with Sensory tasks. The results indicate that this relationship is statistically significant, F(1,29) = 15.329, p = 0.001. Since this result was significant, a t-test was conducted on the individual tasks within a trial. The t-test results for the Sensor task versus the Vision with Sensory task for Trial one were found to be insignificant, t (29) = -0.534, p = 0.597. This result indicates that the difference between the workload values, as shown in Figure 26, as indistinguishable when participants are unable to directly view the environment and the robot. Similar analysis for the same tasks during Trial two was found to be significant, t (29) = -5.565, p < 0.001. This result indicates that the workload value for Sensor task was significantly less than that of the Vision with Sensory task, implying that when participants can directly view the environment and the robot their workload is less with the Sensor task.

The ANOVA results comparing the Sensor task directly with the Vision with Sensory task across both trials was statistically significant, F(1, 29) = 12.150, p = 0.002. The mean perceived workload for the Sensor task was 44.89 (std. dev. = 16.89). This mean is less than the perceived workload mean for the Vision with Sensory task (mean = 55.165, std. dev. = 16.55). This result is an effect of the change to the Sensor task during the second trial. This result may not hold if the participants had been permitted to directly view the environment and the robot while completing the Vision with Sensory task.

An analysis of the Sensor and the Vision with Sensory tasks in Trial one versus Trial two was found to be statistically significant, F(1,29) = 8.925, p = 0.006. This result
indicates that the perceived workload values were statistically different across the trials. The mean perceived workload for Trial one was 53.51 (std. dev. = 17.44). This mean is higher than the mean for Trial two (mean = 46.545, std. dev. = 16). The average workload in aggregate for both tasks during the second trial was lower. The individual task results indicate based on Table 2; that the value increased slightly for the Vision with Sensory task during the second trial. Again, it was anticipated that the perceived workload values between trials would decrease as the participants gained more experience with the system. The reduction found here is due to the fact that participants during Trial two of the Sensor task was able to directly view the robot. The result was a large decrease in the perceived workload values for the Sensor task, which resulted in a statistically significant result across trials.

This subsection shows that the difference between the workload values across the Sensor task and Vision with Sensory task during Trial one was not significant. This result indicates that the perceived mental workload values did not dramatically changed across tasks during Trial one, and stayed relatively the same. But the results indicate that the difference between the workload values for the Sensor task and Vision with Sensory task during Trial two were significant. In this case, the perceived mental workload value for Sensor task is significantly less than the perceived workload for Vision with Sensory task in Trial two. This result shows that the operators were able to complete the Sensor task more easily than the Vision with Sensory task in Trial two.
Vision vs. Vision with Sensory Task Analysis

The repeated measure ANOVA was conducted to determine if a significant relationship existed across trials and tasks for Vision and Vision with Sensory task for the perceived workload values. The results indicate that this relationship is insignificant, F(1, 29) = 0.001, p = 0.973. Since this result was not significant, a T-test was not conducted on the individual tasks within a trial.

The ANOVA results comparing the Vision task directly with the Vision with Sensory task across both trials was statistically significant, F(1, 29) = 27.432, p < 0.001. The mean perceived workload for the Vision task was 43.95 (std. dev. = 16.76). This mean is less than the perceived workload mean for the Vision with Sensory task (mean = 55.165, std. dev. = 16.55). This result indicates that even after repeated use of the two screens, the participants felt that the Vision-only screen required less workload than the Vision with Sensory Overlay screen. This result was anticipated.

An analysis of the combined Vision and the Vision with Sensory tasks results in Trial one versus Trial two was found to be insignificant, F(1,29) =0.333, p =0.568. This result indicates that the perceived workload values did not dramatically changed between trials and stayed relatively the same. The mean perceived workload for Trial one across both tasks was 48.91 (std. dev. = 15.98). While this mean is less than the mean for Trial two (mean = 50.205, std. dev. = 17.33), this difference is rather small. It was anticipated that the workload values would decrease during Trial two. What was found, although it is only a slight increase, was not expected. This result may be due to the fact that during the second trial, participants were able to directly view the robot’s environment during the Sensor task, which resulted in dramatically reduced their perceived workload for that
task. This change to the Sensor task may have caused participants to rate their perceived workloads for the Vision and Vision with Sensory tasks slightly higher, as the results show. It is not clear from the data that this is the cause of this result.

This subsection shows that the difference between the workload values for Vision task and Vision with Sensory task was statistically significant with the Vision task requiring less workload to complete the tasks. The results found slight increases in the workload values for both screens between trials, but this was found to be insignificant.

Discussion

In general, when participants are required to use all three screens in a remote location, the Vision-only screen requires the least amount of workload. This result verifies the first hypothesis claimed in this work. First hypothesis states that: “The Vision with Sensory Overlay screen and Sensor-only screen induce higher workload levels for the defined user group than the Vision-only screen.” In this same situation, the Sensor-only and Vision with Sensory Overlay screens resulted in almost identical workload.

The results indicating that the perceived workload was significantly lower for the Sensor task over the other two tasks when the participants could directly view the environment and robot are not surprising. In fact, this result would be expected. What cannot be determined by this study is if the above stated hypothesis would be valid when the participants were permitted to directly view the environment for all three tasks. It is anticipated that the results would be similar, but with lower workload results, to those found during the first trial of this evaluation.

The results also indicated a slight increase in perceived workload for the Vision and Vision with Sensory tasks form one trial to the next. This was not anticipated and it is
felt that this is a result of the change to the Sensor task during the second trial. It would be anticipated that the perceived workload values would decrease as participants become more familiar with the system and the interface screens.

Likert Scale Questions Analysis

This thesis section provides information regarding the analysis of the Likert scale questions. The complete data is provided in Appendix C. The post-task questionnaire gathered information related to subjective usability. Participants answered four Likert scale questions that used a 5-point scale. The post-task questionnaire was administered after the completion of each task.

Thirty participants completed the user evaluation; therefore there are 30 data points per task per trial resulting in a total of 60 observations per task. A Friedman Analysis of Variance by Ranks Test (referred to as the Friedman Test in this document) was conducted on this data. Results from this analysis are considered significant if \( p \leq 0.05 \). If this analysis was significant, the data was further analyzed using a Wilcoxon Signed-Rank test with Bonferroni-correlated alpha \( (\alpha = 0.018) \). Results of this analysis are considered significant if \( p \leq 0.018 \).

Ease of Using Screen Question Analysis

The first Likert scale question asked the participants to circle the number that most accurately reflects his or her response regarding the ease of using the screen to complete the task. The endpoints of the scale were Very Difficult (1) and Very Easy (5).
Overall Analysis

The descriptive statistics related to the responses for the ease of using a particular screen to complete a task are provided in Table 3. The table indicates that the Vision-only screen mean response increased by 0.50 between the first and second trials. The mean related to the Vision with Sensory Overlay screen increased by 0.34 between trials. The mean response for the Sensor-only screen increased dramatically by 1.60 between trials. This large increase was the result of the participant’s environmental condition changing between the trials as described in Chapter 5 in page 47. The means for all screens went up during the second trial, as shown from the Figure 27.

Table 3: The descriptive statistics related to the perceived ease of using a particular screen

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td>Mean</td>
<td>3.83</td>
<td>2.77</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.83</td>
<td>1.36</td>
</tr>
<tr>
<td><strong>Trial 2</strong></td>
<td>Mean</td>
<td>4.33</td>
<td>4.37</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.80</td>
<td>0.96</td>
</tr>
</tbody>
</table>

The participants gave the highest rating, 4.37 out of 5, to the Sensor-only screen during Trial two; additionally the Sensor-only screen had the lowest rating, 2.77, during Trial one. There was a large change between the trials with this screen caused by the environmental change during Trial two. The Vision-only screen had the highest overall rating across both trials (Vision-only Screen = 4.08, Sensor-only Screen = 3.57, Vision with Sensory Overlay Screen = 3.10). Across both trials Vision with Sensory Overlay
screen had the lowest overall rating. The changes between the trials are shown visually in Figure 27.

![Figure 27: Ease of Use Likert Scale Question Responses for All Screens](image)

**Trial One Analysis**

During Trial one, the participants rated the Vision-only screen as the highest and the Sensor-only screen as the lowest. They rated the Vision-only screen as 3.83, the Sensor-only screen as 2.77, and the Vision with Sensory Overlay screen as 2.93.

The Friedman Test was conducted to identify significant results. The results indicate that there is a significant difference between the responses during Trial one, \( \chi^2(2, N=30) = 14.102, p = 0.001 \). Since this result was significant, further analysis was
conducted to compare individual screen pairs. The Wilcoxon-Signed Rank Test was conducted on the individual screens.

The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial one was found to be statistically significant, $N = 30$, $Z = -3.299$, $p = 0.001$. This result indicates that the participants found the Vision-only screen significantly easier to use than the Sensor-only screen.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial one was found to be insignificant, $N = 30$, $Z = -0.428$, $p = 0.669$. The difference between the means for these two screens was 0.16. This result indicates that neither screen was rated statistically easier to use than the other.

The Wilcoxon results for the Vision-only versus the Vision with Sensory Overlay screens was found to be statistically significant, $N = 30$, $Z = -3.425$, $p = 0.001$. This result indicates that the participants felt that the Vision-only screen is significantly easier to use than the Vision with Sensory Overlay screen.

**Trial Two Analysis**

During Trial two, the participants ranked the Sensor-only screen as the highest and the Vision with Sensory Overlay screen as the lowest. They rated the Vision-only screen as 4.33, the Sensor-only screen as 4.37 and gave a value of 3.27 to the Vision with Sensory Overlay screen.

The Friedman test result indicated that there is a significant difference across screens during Trial two, $\chi^2 (2, N=30) = 18.829$, $p < 0.001$. Since this result was
significant, further analysis was conducted. A Wilcoxon-Signed Rank Test was conducted on the individual interface screen pairs.

The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial two was found to be insignificant, $N = 30, Z = -0.164, p = 0.87$. The difference between the means for these two screens was 0.04. This result indicates that neither screen was rated statistically easier to use than the other. This result also implies that both screens were ranked as easy to use. This result related to the Sensor-only screen is primarily due to the environmental change for the Sensor task in which the participants were permitted to see the robot and its environment during the second trial.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial two was statically significant, $N = 30, Z = -3.364, p = 0.001$. This result indicates that when participants are able to directly view the environment and robot while using the Sensor-only screen, they find this screen to be significantly easier to use than the Vision with Sensory Overlay screen.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens during Trial two was found to be statistically significant, $N = 30, Z = -3.604, p < 0.001$. This result indicates that it was believed that the Vision-only screen was significantly easier to use than the Vision with Sensory Overlay screen.

This subsection shows that during Trial one, the Vision-only screen was found to be the significantly easier to use than the Sensor-only screen. This subsection also found that the Vision-only screen was significantly easier to use than the Vision with Sensory Overlay screen during both trials. It was found that during Trial two, there was no statistical difference in the ratings for the Vision-only and Sensor-only screens. Finally, it
was found that the Sensor-only screen was significantly easier to use than the Vision with Sensory Overlay screen when participants were able to directly view the environment when completing the Sensor task.

Ability to Interpret and Understand Presented Information Question Analysis

The second Likert scale question asked the participants to circle the number that most accurately reflected his or her response regarding the ability to interpret and understand the presented information during the task. The scale endpoints were Rarely (1) and Always (5).

Overall Analysis

The descriptive statistics related to the responses for the interpretation and understanding the presented information during the task are provided in Table 4. The table indicates that the Vision-only screen mean response increased slightly by 0.07 between the first and second trials. The Vision with Sensory Overlay screen mean decreased by 0.10 between trials. The mean response for the Sensor-only screen increased by 1.00 between trials. This increase was the result of the participant’s environmental condition changing between the trials as mentioned before.
Table 4: The descriptive statistics related to the interpretation and understanding the presented information during the task.

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.43</td>
<td>3.57</td>
<td>3.97</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.86</td>
<td>1.16</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Trial 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.50</td>
<td>4.57</td>
<td>3.87</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.68</td>
<td>0.63</td>
<td>0.97</td>
</tr>
</tbody>
</table>

The participants gave the highest rating, 4.57 out of 5, to the *Sensor-only* screen during Trial two; additionally the *Sensor-only* screen had the lowest rating, 3.57 during Trial one. There was a large change between the trials for this screen caused by the environmental change during Trial two. The *Vision-only* screen had the highest overall rating across both trials (*Vision-only* Screen = 4.47, *Sensor-only* Screen = 4.07, *Vision with Sensory Overlay* Screen = 3.92). Across both trials *Vision with Sensory Overlay* screen had the lowest overall rating. The changes between the trials are shown visually in Figure 28.
Figure 28: Ability to Interpret and Understand the Presented Information during the task Responses for All Screens

**Trial One Analysis**

During Trial one, the participants rated the Vision-only screen as the highest and the Sensor-only screen as the lowest. They rated the Vision-only screen as 4.43, the Sensor-only screen as 3.57, and the Vision with Sensory Overlay screen as 3.97.

The Friedman Test was conducted to identify significant results. The results indicate that there is a significant difference between the responses during Trial one, \( \chi^2 (2, N=30) = 12.025, p = 0.002 \). Since this result was significant, further analysis was conducted to compare individual screen pairs using the Wilcoxon-Signed Rank Test.

The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial one was found to be statistically significant, \( N = 30, Z = -3.014, p = 0.003 \).
This result indicates that the participants found the *Vision-only* screen significantly increased their ability to interpret and understand the presented information during the task over the *Sensor-only* screen.

The Wilcoxon test results for the *Sensor-only* versus the *Vision with Sensory Overlay* screens during Trial one was found to be insignificant, \( N = 30, Z = -1.507, p = 0.132 \). The difference between the means for these two screens was 0.40. This result indicates that neither screen was rated statistically better for interpreting and understanding the presented information.

The Wilcoxon test results for the *Vision-only* versus the *Vision with Sensory Overlay* screens during Trial one was found to be insignificant, \( N = 30, Z = -2.118, p = 0.034 \). The difference between the means for these two screens was 0.46. This result indicates that neither screen was rated statistically better for interpreting and understanding the presented information.

**Trial Two Analysis**

During Trial two, the participants ranked the *Sensor-only* screen as the highest and the *Vision with Sensory Overlay* screen as the lowest. They rated *Vision-only* screen as 4.50, the *Sensor-only* screen as 4.57, and gave a value of 3.87 to the *Vision with Sensory Overlay* screen.

The Friedman test results indicate that there is a significant difference during Trial two, \( \chi^2 (2, N=30) = 12.026, p < 0.002 \). Since this result was significant, further analysis was conducted to compare individual screen pairs. The Wilcoxon-Signed Rank Test was conducted.
The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial two was found to be insignificant, N = 30, Z = -0.535, p = 0.593. The difference between the means for these two screens was 0.04. This result indicates that neither screen was rated statistically better with regard to interpreting and understanding the presented information. This result is primarily due to the environmental change for the Sensor task in which the participants were permitted to see the robot and its environment during the second trial.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial two was found to be statically significant, N = 30, Z = -2.992, p = 0.003. This result indicates that the Sensor-only screen was significantly better than the Vision with Sensory Overlay screen with regard to the ability to interpret and understand the presented information.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens during Trial two was found to be statistically significant, N = 30, Z = -3.042, p = 0.002. This result indicates that the Vision-only screen was significantly better than the Vision with Sensory Overlay screen with regard to the ability to interpret and understand the presented information.

This subsection shows that during Trial one, the participants interpreted and understood the presented information during the task significantly better with the Vision-only screen than the Sensor-only screen. This subsection found that the participants interpreted and understood the presented information during the task significantly better with the Sensor-only screen than with the Vision with Sensory Overlay screen during Trial two. It was also found that the participants understood the presented information
better with the *Vision-only* screen than the *Vision with Sensory Overlay* screen during the second trial.

**Feeling in Control of the Robot Likert Scale Question Analysis**

The third Likert scale question asked the participants to circle the number that most accurately reflects his or her response regarding their ability to feel in control of the robot during the task. The endpoints of the scale were *Rarely* (1) and *Always* (5).

**Overall Analysis**

The descriptive statistics related to the responses for the feeling in control of the robot during the task are provided in Table 5. The table indicates that the *Vision-only* screen mean response increased by 0.33 between the first and second trials. The *Vision with Sensory Overlay* screen mean reflects a 0.26 increase between trials. The mean response for the *Sensor-only* screen increased by 0.90 between trials. This increase was the result of the participant’s environmental condition changing between the trials. The means for all screens went up in the second trial as seen from the Figure 29.

**Table 5:** The descriptive statistics related to feeling in control of the robot during the task

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.80</td>
<td>3.20</td>
<td>2.77</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.13</td>
<td>1.35</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Trial 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.13</td>
<td>4.10</td>
<td>3.03</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.00</td>
<td>1.18</td>
<td>1.03</td>
</tr>
</tbody>
</table>
The participants gave the highest rating, 4.13 out of 5, to the Vision-only screen during Trial two; and the Vision with Sensory Overlay screen had the lowest rating, 2.77, during Trial one. The Vision-only screen had the highest overall rating across both trials (Vision-only Screen = 3.97, Sensor-only Screen = 3.65, Vision with Sensory Overlay Screen = 2.90). Across both trials the Vision with Sensory Overlay screen had the lowest overall rating. The changes between the trials are shown visually in Figure 29.

Figure 29: Feeling in Control of the Robot during the task Question Responses for All Screens

**Trial One Analysis**

During Trial one, the participants rated the Vision-only screen as the highest and the Vision with Sensory Overlay screen as the lowest. They rated the Vision-only screen
as 3.80, the *Sensor-only* screen as 3.20, and the *Vision with Sensory Overlay* screen as 2.77.

The Friedman Test was conducted to identify significant results. The results indicate that there is a significant difference between the responses during Trial one, \( \chi^2(2, N=30) = 11.113, p = 0.004 \). Since this result was significant, further analysis was conducted to compare individual screen pairs. Wilcoxon-Signed Rank Test was conducted on the individual interface screens.

The Wilcoxon test results for the *Vision-only* versus the *Sensor-only* screens during Trial one was found to be insignificant, \( N = 30, Z = -1.936, p = 0.053 \). The difference between the means for these two screens was 0.60. This result indicates that neither screen was rated as providing statistically better control of the robot during the task over the other.

The Wilcoxon test results for the *Sensor-only* versus the *Vision with Sensory Overlay* screens during Trial one was found to be insignificant, \( N = 30, Z = -1.439, p = 0.150 \). The difference between the means for these two screens was 0.43. This result indicates that neither screen was rated statistically better with regard to feeling in control of the robot during the task than the other.

The Wilcoxon test results for the *Vision-only* versus the *Vision with Sensory Overlay* screens during Trial one was found to be statistically significant, \( N = 30, Z = -3.624, p < 0.001 \). This result indicates that the participants felt that the *Vision-only* screen provided a significantly better feeling of control over the robot than the *Vision with Sensory Overlay* screen.
Trial Two Analysis

During Trial two, the participants ranked the Vision-only screen as the highest and the Vision with Sensory Overlay screen as the lowest. They rated the Vision-only screen as 4.13, the Sensor-only screen as 4.10, and gave a value of 3.03 to the Vision with Sensory Overlay screen.

The Friedman test result indicates that there is a significant difference during Trial two, $\chi^2 (2, N=30) = 22.615, p < 0.001$. Since this result was significant, further analysis was conducted to compare individual screen pairs. A Wilcoxon-Signed Rank Test was conducted on the individual interface screen pairs.

The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial two was found to be insignificant, $N = 30, Z = -0.175, p = 0.861$. The difference between the means for these two screens was 0.03. This result indicates that neither screen was rated statistically better with regard to feeling in control of the robot during the task over the other.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial two was statically significant, $N = 30, Z = -3.795, p < 0.001$. This result indicates that when participants are able to directly view the robot and its environment while using the Sensor-only screen, they find this screen to provide a significantly better feeling of control over the robot than the Vision with Sensory Overlay screen.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens during Trial two was found to be statistically significant, $N = 30, Z =$
-3.835, p < 0.001. This result indicates that the participants found the Vision-only screen to provide significantly better control over the robot than the Vision with Sensory Overlay screen.

This subsection shows that during both trials, the participants felt significantly more in control of the robot with the Vision-only screen than with the Vision with Sensory Overlay screen. This subsection also found that during Trial two, the participants felt significantly more control of the robot with the Sensor-only screen than the Vision with Sensory Overlay screen.

Ability to Correct Errors Likert Scale Question Analysis

The fourth Likert scale question asked the participants to circle the number that most accurately reflects his or her response regarding the ability to correct errors during the task. The scale endpoints were Rarely (1) and Always (5).

Overall Analysis

The descriptive statistics related to the responses for the ability to correct errors during the task are provided in Table 6. The table indicates that the Vision-only screen mean response increased by 0.10 between the first and second trials. The Vision with Sensory Overlay screen mean increased by 0.23 between trials. The mean response for the Sensor-only screen increased dramatically by 1.43 between trials. This large increase was the result of the participant’s environmental condition changing between the trials as described in Chapter 5 in page 47. The means for all screens went up in the second trial as seen from the Figure 30.
Table 6: The descriptive statistics related to the ability to correct errors during the task

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td><strong>Mean</strong></td>
<td>4.30</td>
<td>3.47</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.84</td>
<td>1.31</td>
<td>1.11</td>
</tr>
<tr>
<td><strong>Trial 2</strong></td>
<td><strong>Mean</strong></td>
<td>4.40</td>
<td>4.40</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.72</td>
<td>0.77</td>
<td>0.99</td>
</tr>
</tbody>
</table>

The participants gave the highest rating, 4.40 out of 5, to the Vision-only and Sensor-only screens during Trial two. The Vision with Sensory Overlay screen had the lowest rating, 3.07, during Trial one. The Vision-only screen had the highest overall rating across both trials (Vision-only Screen = 4.35, Sensor-only Screen = 3.94, Vision with Sensory Overlay Screen = 3.19). Across both trials, Vision with Sensory Overlay screen had the lowest overall rating. The changes between the trials are shown visually in Figure 30.
During Trial one, the participants rated the Vision-only screen as the highest and the Vision with Sensory Overlay screen as the lowest. They rated the Vision-only screen as 4.30, the Sensor-only screen as 3.47, and the Vision with Sensory Overlay screen as 3.07.

The Friedman was conducted to identify significant results. The results indicate that there is a significant difference between responses during Trial one, \( \chi^2 (2, N=30) = 14.866, p = 0.001 \). Since this result was significant, further analysis was conducted to compare individual screen pairs. A Wilcoxon-Signed Rank Test was conducted.
The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial one was found to be statistically significant, \( N = 30, Z = -2.419, p = 0.016 \). This result indicates that the participants found the Vision-only screen significantly better with regard to their ability to correct errors during the task than the Sensor-only screen.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial one was found to be insignificant, \( N = 30, Z = -1.413, p = 0.158 \). The difference between the means for these two screens was 0.40. This result indicates that neither screen was rated statistically better with regard correcting errors during the task than the other.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens was found to be statistically significant, \( N = 30, Z = -3.898, p < 0.001 \). This result indicates that the participants felt the Vision-only screen is significantly better with regard to their ability to correct errors during the task than the Vision with Sensory Overlay screen.

**Trial Two Analysis**

During Trial two, the participants ranked the Vision-only and Sensor-only screens as the highest and the Vision with Sensory Overlay screen as the lowest. They rated Vision-only screen as 4.40, the Sensor-only screen as 4.40, and gave a value of 3.30 to the Vision with Sensory Overlay screen.

The Friedman test result indicates that there is a significant difference during Trial two, \( \chi^2 (2, N=30) = 20.725, p < 0.001 \). Since this result was significant, further analysis was conducted to compare individual screen pairs using a Wilcoxon-Signed Rank Test.
The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial two was found to be insignificant, N = 30, Z = -0.133, p = 0.894. The difference between the means for these two screens was 0.00. This result indicates that neither screen was rated statistically better with regard to their ability to correct errors during the task than the other. This result is primarily due to the environmental change for the Sensor task in which the participants were permitted to see the robot and its environment during the second trial.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial two was statically significant, N = 30, Z = -3.814, p < 0.001. This result indicates that when participants are able to directly view the environment and robot while using the Sensor-only screen, they find this screen to be significantly better with regard to correcting errors during the task than the Vision with Sensory Overlay screen.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens during Trial two was found to be statistically significant, N = 30, Z = -3.814, p < 0.001. This result indicates that the Vision-only screen was found to be significantly better for correcting errors during the task than the Vision with Sensory Overlay screen.

This subsection shows that during Trial one, the Vision-only screen was found to be significantly better with regard to the ability to correct errors during the task than the Sensor-only screen. This subsection also found that the Vision-only screen was significantly better than the Vision with Sensory Overlay screen with regard to the ability to correct errors during the task during both trials. It was also found that during Trial two,
the participants felt significantly higher ability to correct their errors during the task with the Sensor-only and Vision-only screens than the Vision with Sensory Overlay screen.

Discussion

When the participants are required to use all three screens from a remote location, they rated the Vision-only screen as significantly easier to use than the Sensor-only and the Vision with Sensory Overlay screens.

When participants' environmental condition changed during Trial two of the Sensor task, the previous result no longer holds. This would be expected, since during Trial two of the Sensor-only task, the participants were permitted to see the robot and its environment, therefore the Sensor-only screen was ranked as easiest to use. During Trial two it was found that Vision-only screen was significantly easier to use than the Vision with Sensory Overlay screen. It was also found that when the participants are required to use all three screens in a remote location, the Vision-only screen provides a significantly better ability to interpret and understand the presented information during the task than the Sensor-only screen. This result shows that the visual information provides a better means of understanding the remote environment as the vision-based screen received a higher rating than the sensor-based screen.

When participants' environmental condition changed during Trial two, the previous result did not hold. It is found that the participants interpreted and understood the presented information during the task significantly better with the Sensor-only screen than the Vision with Sensory Overlay screen. In fact this result would be expected, since with the Sensor-only screen the participants can see the robot and its environment. It was also found that the participants understood the presented information significantly better
with the Vision-only screen than the Vision with Sensory Overlay screen. This result did not change, since the environmental change did not affect the tasks of the Vision-only and Vision with Sensory Overlay screens.

The analysis showed that the participants felt significantly more in control of the robot while using the Vision-only screen than when using the Vision with Sensory Overlay screen during both trials. This analysis also showed that the participants felt significantly more in control of the robot with the Sensor-only screen over the Vision with Sensory Overlay screen during the second trial. It is also interesting to note that there is no dramatic difference between the results for the Vision-only screen and the Sensor-only screens when the participants were allowed to see the robot and its environment during the Sensor task.

Finally, participants found it significantly easier to correct their errors with the Vision-only screen than with the Sensor-only and the Vision with Sensory Overlay screens when all tasks were executed from a remote environment. When the condition changed, the result changed. The participants felt significantly better at correcting their errors with the Sensor-only screen over the Vision with Sensory Overlay screen; but the second result did not changed.

The combined Likert scale results indicate that when the participants are required to use all three screens from a remote location, they chose the Vision-only screen as significantly easier to use than the Sensor-only and the Vision with Sensory Overlay screens. This result provides partial verification of hypothesis two, which states: “The Image-only screen is easier to use for the defined user group than the Image with Sensory Overlay screen and Sensor-only screen.” The results during the second trial also
supported this hypothesis for the Vision-only screen over the Vision with Sensory Overlay screen. As this hypothesis was also partially verified from the Ranking question responses. Therefore, this provides the verification of the second hypothesis.

**Ranking Questions Analysis**

This section provides information regarding the analysis of the ranking questions. The post-trial questionnaire had the participants’ rank screens along five factors. The post-trial questionnaire was administered after the completion of the first trial of all three tasks and after the second trial of all three tasks. Participants ranked each screen as: 1 – lowest rank, 2 – middle rank or 3 – highest rank. Each value could be used only once for each question. The complete data is provided in Appendix D.

A Friedman Analysis of Variance by Ranks Test (referred to as the Friedman Test in this document) was conducted on this data. Results from this analysis are considered significant if p <= 0.05. If this result was significant, the data was further analyzed using Wilcoxon Signed-Rank test with Bonferroni-correlated alpha (alpha = 0.018). Results of this analysis are considered significant if p <= 0.018.

**Ease of Use Ranking Question Analysis**

The first ranking question asked the participants to rank the three screens according to the perceived ease of use.
Overall Analysis

The descriptive statistics for the Ease of use ranking question responses were calculated for all screens and are provided in Table 7. The table indicates that the mean response for the Vision-only screen decreased by 0.34 between the first and second trials. The mean for the Vision with Sensory Overlay screen decreased by 0.17 between trials, while the mean for the Sensor-only screen increased by 0.50 between trials. This increase was the result of the participants’ environmental condition changing between the trials as described in the Task Descriptions subsection of Chapter 5 in page 47.

Table 7: The descriptive statistics for the ranking question related to the ease of use

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td>2.57</td>
<td>1.73</td>
<td>1.70</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.63</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Trial 2</strong></td>
<td>2.23</td>
<td>2.23</td>
<td>1.53</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.57</td>
<td>0.86</td>
<td>0.82</td>
</tr>
</tbody>
</table>

The participants gave the highest ranking, 2.57 out of 3, to the Vision-only screen during Trial one; moreover the Vision-only screen had the highest overall ranking across both trials (Vision-only Screen = 2.4, Sensor-only Screen = 1.98, Vision with Sensory Overlay screen = 1.62). The participants gave the lowest ranking, 1.53 out of 3, to the Vision with Sensory Overlay screen during Trial two. Across both trials, the Vision with Sensory Overlay screen had the lowest overall ranking. There was a large change between trials of the Sensor-only screen. This is caused by environmental change during
Trial two for the *Sensor* task. The changes between the trials are shown visually in Figure 31.

![Figure 31: Ease of Use Ranking Question Responses for All Screens](image)

**Trial One Analysis**

During Trial one, the participants ranked the *Vision-only* screen as the highest and the *Vision with Sensory Overlay* screen as the lowest. They ranked the *Vision-only* screen as 2.57, the *Sensor-only* screen as 1.73, and the *Vision with Sensory Overlay* screen as 1.70.

The Friedman test was conducted to identify significant results across the three screens. The results indicate that there is a significant difference in the rankings of the screens during Trial one, $\chi^2 (2, N=30) = 14.47, p = 0.001$. Since this result was
significant, further analysis was conducted to compare individual screen pairs. The Wilcoxon-Signed Rank Test was conducted on the individual interface screens.

The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial one was found to be statistically significant at the $p \leq 0.018$ level of significance, $N = 30$, $Z = -3.26$, $p = 0.001$. This result indicates that the participants ranked the Vision-only screen as significantly easier to use than the Sensor-only screen.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial one was found to be insignificant, $N = 30$, $Z = -0.075$, $p = 0.940$. The difference between the means for these two screens was 0.33. This result indicates that neither screen was ranked statistically easier to use than the other.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens during Trial one was found to be statistically significant, $N = 30$, $Z = -3.214$, $p = 0.001$. This result indicates that the Vision-only screen was ranked as significantly easier to use than the Vision with Sensory Overlay screen.

**Trial Two Analysis**

During Trial two, the participants ranked the Vision-only and Sensor-only screens as the highest and the Vision with Sensory Overlay screen as the lowest. They ranked both Vision-only and Sensor-only screens as 2.23 and gave a value of 1.53 to the Vision with Sensory Overlay screen.

The Friedman Test results indicate that there is a significant difference in the rankings of the three screens during Trial two, $\chi^2(2, N=30) = 9.80$, $p = 0.007$. Since this result was significant, further analysis was conducted to compare individual screen pairs. The Wilcoxon-Signed Rank Test was conducted on the individual interface screens.
The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial two was found to be insignificant, N = 30, Z = -0.034, p = 0.973. The mean rankings for these two screens were identical during Trial two. Therefore, this result implies that both screens were ranked as easy to use. This result is primarily due to the environmental change for the Sensor task in which the participants were permitted to see the robot and its environment during the second trial.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial two was found to be insignificant at the p <= 0.018 level, N = 30, Z = -2.210, p = 0.027. The difference between the two means was 0.7. This result implies that there is little difference in the ranking results for these two screens with regard to ease of use.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens during Trial two was found to be statistically significant, N = 30, Z = -2.911, p = 0.004. This result indicates that the Vision-only screen is significantly easier to use than the Vision with Sensory Overlay screen during Trial two.

This subsection shows that during Trial one, the participants ranked the Vision-only screen as significantly easier to use than the Sensor-only and the Vision with Sensory Overlay screens. This subsection shows that during Trial two, the Vision-only screen was significantly easier to use than the Vision with Sensory Overlay screen. The Vision-only and the Sensor-only screens were similarly ranked during Trial two; this was due to the participant’s ability to directly view the environment and the robot during the second trial of the Sensor task.
Ability to Understand Presented Data Ranking Question Analysis

The second ranking question asked the participants to rank the three screens according to their ability to understand the data from the robot on each interface screen.

*Overall Analysis*

The descriptive statistics for the ranking question responses related to the participants’ ability to understand the data from the robot for all screens are presented in Table 8. The table indicates that the mean response for the Vision-only screen decreased by 0.17 between trials. The Vision with Sensory Overlay screen response decreased between trials by 0.10. The mean for the Sensor-only screen increased by 0.26 between trials. This increase was the result of the participants’ environmental condition changing between the trials as previously mentioned.

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td>Mean</td>
<td>2.40</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.72</td>
<td>0.82</td>
</tr>
<tr>
<td><strong>Trial 2</strong></td>
<td>Mean</td>
<td>2.23</td>
<td>2.03</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.68</td>
<td>0.89</td>
</tr>
</tbody>
</table>

The participants gave the highest ranking, 2.40 out of 3, to the Vision-only screen during Trial one; moreover the Vision-only screen had the highest overall ranking, across both trials (Vision-only Screen = 2.32, Sensor-only Screen = 1.9, Vision with Sensory Overlay screen = 1.78). The participants gave the lowest ranking, 1.73 out of 3, to the
Vision with Sensory Overlay screen during Trial two. Across both trials, the Vision with Sensory Overlay screen had the lowest overall ranking. There was a large change between trials for Sensor-only screen. This is caused by environmental change during Trial two for the Sensor task. The changes between the trials are shown visually in Figure 32.

![Figure 32: Ability to Understand Presented Data Ranking Question Response for All Screens](image)

**Trial One Analysis**

During Trial one, the participants ranked the Vision-only screen as highest and the Sensor-only screen as lowest. They ranked the Vision-only screen as 2.40, the Sensor-only screen as 1.77, and the Vision with Sensory Overlay screen as 1.83.
The Friedman Test was conducted on the rankings to identify significant results. The results indicate that there was a significant difference in the screen rankings with regard to the participants’ ability to understand the data presented from the robot during Trial one, \( \chi^2 (2, N=30) = 7.27, p = 0.026 \). Since this result was significant, further analysis was conducted to compare individual screen pairs employing the Wilcoxon-Signed Rank Test across the individual interface screens.

The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial one was found to be insignificant, \( N = 30, Z = -2.295, p = 0.022 \). The difference between the two means was 0.63. This result indicates that neither screen was ranked statistically higher than the other with regard to the participants’ ability to understand the data presented from the robot.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial one was found to be insignificant, \( N = 30, Z = -0.321, p = 0.748 \). The difference between the two means was 0.06. This result implies that there is little difference in the ranking results for these two screens with regard to the participants’ ability to understand the data presented from the robot.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens during Trial one was also found to be insignificant, \( N = 30, Z = -2.315, p = 0.021 \). The difference between the means was 0.57. This result indicates that neither screen was ranked statistically better than the other with regard to the participants’ ability to understand the data presented from the robot.
**Trial Two Analysis**

During Trial two, the participants ranked the *Vision-only* screen as the highest and the *Vision with Sensory Overlay* screen as the lowest. They ranked the *Vision-only* screen as 2.23, the *Sensor-only* screen as 2.03, and the *Vision with Sensory Overlay* screen as 1.73.

The Friedman Test results during Trial two indicated that there was no significant difference in the rankings related to the participants’ ability to understand the information from the robot, $\chi^2 (2, N=30) = 3.80, p = 0.15$. Since this result was insignificant, further analysis was not conducted to compare individual screen pairs.

This subsection shows that there were no significant differences among the means for all of the screens in relation to ranking the screens based upon the ability to understand the presented information.

### Ability to Understand the Robot’s Activities Ranking Question Analysis

The third ranking question asked the participants to rank the three screens according to their ability to understand what the robot was doing.

### Overall Analysis

The descriptive statistics related to the participants’ ability to understand the robot’s activities are provided in Table 9. The table indicates that the mean response for the *Vision-only* screen decreased by 0.17 between the first and second trials. The *Vision with Sensory Overlay* screen mean decreased slightly by 0.03 between trials. The *Sensor-
only screen mean increased by 0.20 between trials. This increase was the result of the participant’s environmental condition changing between the trials.

Table 9: The descriptive statistics for the ranking question related to the participants’ ability to understand the robot’s activities

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>Mean</td>
<td>2.30</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.75</td>
<td>0.85</td>
</tr>
<tr>
<td>Trial 2</td>
<td>Mean</td>
<td>2.13</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.63</td>
<td>0.95</td>
</tr>
</tbody>
</table>

The participants gave the highest ranking, 2.30 out of 3, to the Vision-only screen during Trial one; moreover the Vision-only screen had the highest overall ranking, across both trials (Vision-only Screen = 2.22, Sensor-only Screen = 1.90, Vision with Sensory Overlay screen = 1.89). The participants gave the lowest ranking, 1.80 out of 3, to the Sensor-only screen during Trial one. Across both trials, the Vision with Sensory Overlay screen had the lowest overall ranking. There was a large change between trials for Sensor-only screen. This is caused by environmental change during Trial two for the Sensor task. The changes between the trials are shown visually in Figure 33.
Figure 33: Ability to Understand the Robot's Activities Ranking Question Response for All Screens

**Trial One Analysis**

During Trial one, the participants ranked the *Vision-only* screen as the highest and the *Sensor-only* screen as the lowest. They ranked the *Vision-only* screen as 2.30, the *Sensor-only* screen as 1.80, and the *Vision with Sensory Overlay* screen as 1.90.

The Friedman Test was conducted on the ranking data to identify significant results. The results indicate that there is no significant difference in the rankings of the three screens during Trial one, $\chi^2 (2, \text{N}=30) = 4.20, p = 0.122$. Since this result was insignificant, further analysis was not conducted to compare individual screen pairs. The mean difference between the *Vision-only* and *Sensor-only* screens was 0.5. The mean difference between the *Vision-only* and the *Vision with Sensory Overlay* screens was 0.4
while the mean difference between the Sensor-only and the Vision with Sensory Overlay screens was 0.1.

*Trial Two Analysis*

During Trial two, the participants ranked the Vision-only screen as the highest and the Vision with Sensory Overlay screen as the lowest. They ranked the Vision-only screen as 2.13, the Sensor-only screen as 2.00, and the Vision with Sensory Overlay screen as 1.87.

The Friedman Test results indicate that there was no significant difference in the rankings of the three screens during Trial two, $\chi^2 (2, N=30) = 1.07, p = 0.587$. Therefore, further analysis was not conducted to compare individual screen pairs. The mean difference between the Vision-only and Sensor-only screens was 0.13. The mean difference between the Vision-only and the Vision with Sensory Overlay screens was 0.26 while the mean difference between the Sensor-only and the Vision with Sensory Overlay screens was 0.13.

This subsection shows that there was no significant difference among the means of all the screens related to the participants’ ability to understand what the robot is doing.

**Ability to Understand the Robot’s Working Environment Ranking Question Analysis**

The fourth ranking question asked the participants to rank the three screens according to their ability to understand the environment in which the robot was working.
Overall Analysis

The descriptive statistics for the question response for all screens related to the participants’ ability to understand the robot’s working environment are provided in Table 10. The table indicates that the Vision-only screen mean response decreased slightly by 0.07 between the first and second trials. The Vision with Sensory Overlay screen mean decreased by 0.46 between trials. Finally, the Sensor-only screen mean increased by 0.53 between trials. This increase was the result of the participants’ environmental condition changing between the trials as described previously.

Table 10: The descriptive statistics for the ranking question related to the participants’ ability to understand the robot’s working environment

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.10</td>
<td>1.57</td>
<td>2.33</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.76</td>
<td>0.77</td>
<td>0.76</td>
</tr>
<tr>
<td><strong>Trial 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.03</td>
<td>2.10</td>
<td>1.87</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.67</td>
<td>0.88</td>
<td>0.90</td>
</tr>
</tbody>
</table>

The participants gave the highest ranking, 2.33 out of 3, to the Vision with Sensory Overlay screen during Trial one. They provided the highest overall ranking, across both trials to the Vision with Sensory Overlay screen (Vision-only Screen = 2.07, Sensor-only Screen = 1.84, Vision with Sensory Overlay screen = 2.10). The participants gave the lowest ranking, 1.57 out of 3, to the Sensor-only screen during Trial one. Across both trials, the Sensor-only screen had the lowest overall ranking. The mean changes between the trials are shown visually in Figure 34.
During Trial one, the participants ranked the Vision with Sensory Overlay screen as the highest and the Sensor-only screen as the lowest. They ranked the Vision-only screen as 2.10, the Sensor-only screen as 1.57, and the Vision with Sensory Overlay screen as 2.33.

The Friedman Test was conducted to identify significant results. The results indicate that there was a significant difference in the rankings of the three screens during Trial one, $\chi^2 (2, N=30) = 9.27, p = 0.01$. Since this result was significant, further analysis was conducted to compare individual screen pairs. The Wilcoxon-Signed Rank Test was conducted to compare the individual screen pairs.
The Wilcoxon test results for the Vision-only versus the Sensor-only screens during Trial one was found to be insignificant, N = 30, Z = -1.993, p = 0.046. The difference between the means is 0.53. This result indicates that neither screen was ranked statistically better with regard to the participants’ ability to understand the robot’s working environment.

The Wilcoxon test for the Sensor-only versus the Vision with Sensory Overlay screens during Trial one was found to be statistically significant, N = 30, Z = -2.772, p = 0.006. This result indicates that the participants felt that while located in a remote environment, the Vision with Sensory Overlay screen significantly improved their ability to understand the robot’s working environment over the Sensor-only screen.

The Wilcoxon test results for the Vision-only versus the Vision with Sensory Overlay screens during Trial one was found to be insignificant, N = 30, Z = -0.948, p = 0.343. The difference between the means is 0.23. This result indicates that neither screen was ranked statistically better than the other with regard to the participants’ ability to understand the robot’s working environment.

**Trial Two Analysis**

During Trial two, the participants ranked the Sensor-only screen as highest and the Vision with Sensory Overlay screen as lowest. They ranked the Vision-only screen as 2.03, the Sensor-only screen as 2.10, and the Vision with Sensory Overlay screen as 1.87.

The Friedman Test result for the second trial indicates that there was an insignificant relationship between the three screens, \( \chi^2 (2, N=30) = 0.87, p = 0.648 \). Therefore, further analysis was not conducted to compare individual screen pairs. The mean difference between the Vision-only screen and Sensor-only screen was 0.07, and the
difference of the mean between the Vision-only and Vision with Sensory Overlay screen was 0.16. Finally, the mean difference between the Sensor-only and Vision with Sensory Overlay screens is 0.23.

This subsection shows that the Vision with Sensory Overlay screen provided the ability to better understand the environment in which the robot is working over the Sensor-only screen during Trial one. This result is related to the information provided in these screens. The visual information is the primary difference between the two screen displays. The sensory information was provided in both screens, but the vision information was provided with the Vision with Sensory Overlay screen but not with the Sensor-only screen. This result indicates that the visual information provides a better understanding of the environment in which the robot is working. It is felt that users are familiar with images and that images provide a large amount of information to the users. Additionally, the Vision with Sensory Overlay screen was ranked higher than the Vision-only screen, which is logical since Vision with Sensory Overlay screen provides both vision and sensory information therefore it provides better understanding of the environment in which the robot is working, but this result was not statistically significant.

General Overall Ranking Question Analysis

The fifth ranking question asked the participants to rank the three screens according to their general overall ranking.

Overall Analysis

The descriptive statistics for the question asking the participants to provide an overall rank of the three screens is provided in Table 11. The table indicates that the
Vision-only screen mean decreased by 0.27 between the first and second trials. The Vision with Sensory Overlay screen mean decreased slightly by 0.04 between trials. Finally, the Sensor-only screen mean between trials increased by 0.30. As previously mentioned, this increase was the result of the participant’s environmental condition changing between the trials.

Table 11: The descriptive statistics for the overall ranking

<table>
<thead>
<tr>
<th></th>
<th>Vision-only Screen</th>
<th>Sensor-only Screen</th>
<th>Vision with Sensory Overlay Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1 Mean</td>
<td>2.50</td>
<td>1.63</td>
<td>1.87</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.63</td>
<td>0.76</td>
<td>0.82</td>
</tr>
<tr>
<td>Trial 2 Mean</td>
<td>2.23</td>
<td>1.93</td>
<td>1.83</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.68</td>
<td>0.91</td>
<td>0.83</td>
</tr>
</tbody>
</table>

The participants gave the highest ranking, 2.50 out of 3, to the Vision-only screen during Trial one; moreover the Vision-only screen had the highest overall ranking, across both trials (Vision-only Screen = 2.37, Sensor-only Screen = 1.78, Vision with Sensory Overlay screen = 1.35). The participants gave the lowest ranking, 1.63 out of 3, to the Sensor-only screen during Trial one. Across both trials, the Vision with Sensory Overlay screen had the lowest overall ranking. There was a large change between trials for Sensor-only screen. This is caused by environmental change during Trial two for the Sensor task. The changes between the trials are shown visually in Figure 35.
Figure 35: General overall Ranking Question Response for All Screens

**Trial One Analysis**

During Trial one, the participants ranked the *Vision-only* screen as the highest and the *Sensor-only* screen as the lowest. They ranked the *Vision-only* screen as 2.50, the *Sensor-only* screen as 1.63, and the *Vision with Sensory Overlay* screen as 1.87.

The Friedman Test was conducted on the participants’ overall ranking data in order to identify significant results. The results indicate that there was a significant difference across the screens during Trial one, $\chi^2 (2, N=30) = 12.07$, $p = 0.002$. Since this result was significant, further analysis was conducted to compare individual screen pairs. Wilcoxon-Signed Rank Test was conducted on the individual interface screens.
The Wilcoxon test comparing the Vision-only versus the Sensor-only screens during Trial one was found to be statistically significant, \( N = 30, Z = -3.215, p = 0.001 \). This result indicates that the overall ranking of the Vision-only screen is significantly higher than the response for the Sensor-only screen.

The Wilcoxon test results for the Sensor-only versus the Vision with Sensory Overlay screens during Trial one was found to be insignificant, \( N = 30, Z = -1.014, p = 0.311 \). The mean difference between the two screens was 0.24. This result indicates that neither screen was ranked statistically better than the other with regard to the participants’ overall ranking.

The Wilcoxon results for the Vision-only versus the Vision with Sensory Overlay screens during Trial one was found to be statistically significant, \( N = 30, Z = -2.638, p = 0.008 \). This result indicates that the mean overall ranking for the Vision-only screen is significantly higher than the mean response for the Vision with Sensory Overlay screen.

**Trial Two Analysis**

During Trial two, the participants ranked the Vision-only screen as highest and the Vision with Sensory Overlay screen as lowest. They ranked the Vision-only screen as 2.23, the Sensor-only screen as 1.93, and the Vision with Sensory Overlay screen as 1.83.

The Friedman Test results indicate that an insignificant relationship existed between the rankings of all three screens during Trial two, \( \chi^2 (2, N=30) = 2.6, p = 0.273 \). Since this result was insignificant, further analysis was not conducted to compare individual screen pairs. The mean difference between the Vision-only and Sensor-only screens is 0.3, between the Vision-only and the Vision with Sensory Overlay screens is
0.4, and finally between the Sensor-only and the Vision with Sensory Overlay screens is 0.1.

The results from this subsection found that participants felt that the Vision-only screen ranked not only the highest of the three screens but significantly higher than the Sensor-only and the Vision-only screens during Trial one. A similar result was not found during Trial two. It is felt that the result changed due to the environmental change with the Sensor task during Trial two.

Discussion

When the participants are required to use all three screens from a remote location, they rated the Vision-only screen as significantly easier to use than the Sensor-only and the Vision with Sensory Overlay screens. This result provides partial verification of hypothesis two, which states: “The Image-only screen is easier to use for the defined user group than the Image with Sensory Overlay screen and Sensor-only screen.”

When the participants’ environmental condition changed during Trial two of the Sensor task, the verification of the hypothesis no longer holds. This would be expected, since during Trial two of the Sensor task, the participants were permitted to see the robot and its environment, therefore the Sensor-only screen was ranked as easiest to use. During Trial two it was found that Vision-only screen was significantly easier to use than the Vision with Sensory Overlay screen, therefore the portion of the hypothesis related to these two screens still holds.

It was also found that when the participants are required to use all three screens from a remote location, the Vision with Sensory Overlay screen provides a significantly better ability to understand the environment in which the robot was working than with the
Sensor-only screen. This result shows that the visual information provides a better means of understanding the remote environment as the vision-based screen received a higher ranking than the sensor-based screen.

Finally, the participants gave a significantly higher general overall ranking to the Vision-only screen than the Sensor-only and Vision with Sensory Overlay screens when the participants were required to use all three screens from a remote location. This provides the verification of the third hypothesis: “The Image-only screen is the preferred teleoperation interface by the defined user group.”

Unfortunately, hypothesis four could not be completely verified through this evaluation. It was claimed that “The Image with Sensory Overlay screen is preferred for teleoperation over the Sensor-only screen by the defined user group.” The results show that during Trial one, participants provided overall rankings of 2.50 for the Vision-only screen, 1.63 for the Sensor-only screen, and 1.87 for the Vision with Sensory Overlay screen. These results were not statistically verified by the analysis. The results changed for Sensor-only and Vision with Sensory Overlay screens during Trial two. The participants provided overall rankings of 2.23 for the Vision-only screen, 1.93 for the Sensor-only screen, and 1.83 for the Vision with Sensory Overlay screen. In fact this result would be expected, since during this trial the participants were able to see the robot and environment while completing the Sensor task. This also shows that the environmental change did not affect the Vision-only screen’s high overall ranking.
Task Completion Times

This section provides information regarding the task completion times for all three tasks. Table 12 shows the descriptive statistics across tasks and trials for the task completion times. The data in the table is provided in seconds. The detailed completion times for all participants are provided in Appendix E.

Table 12: The descriptive statistics related to the Completion Times for All Tasks across Trials

<table>
<thead>
<tr>
<th></th>
<th>Vision Task</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>258.23</td>
<td>215.5</td>
<td>291.00</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>53.51</td>
<td>82.52</td>
<td>22.69</td>
</tr>
<tr>
<td>Trial 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>239.8</td>
<td>111.97</td>
<td>279.33</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>49.38</td>
<td>72.25</td>
<td>34.56</td>
</tr>
</tbody>
</table>

The average completion time for the Vision task during Trial one was 258.23 seconds (approximately 4 minutes 18 seconds). The average during Trial two was 239.8 seconds (approximately 4 minutes). Therefore, the Vision task completion times decreased from Trial one to Trial two.

The average Sensor task completion time during Trial one was 215.5 seconds (approximately 3 minutes 36 seconds). The average Sensor task completion time during Trial two was 111.97 seconds (approximately 1 minute 52 seconds). The completion times decreased dramatically between trials. This decrease occurred because the participants’ were permitted to directly view the environment and robot between trials.

The completion time average for the Vision with Sensory task during Trial one was 291 seconds (4 minutes 51 seconds). The average completion time during Trial two
was 279.33 seconds (approximately 4 minutes and 39 seconds). Therefore, the completion times decreased during the *Vision with Sensory* tasks between trials.

The completion times for all screens across trials are shown visually in Figure 36. According to the figure, the participants completed the task with the *Sensor-only* screen fastest, and the slowest task completion time was found for the *Vision with Sensory* during Trial one. This result is also found during Trial two.

![Figure 36: Completion Times for all screens across trials](image)

The completion time for the *Sensor* task was the shortest compared to the others. This result does not represent that this task was the easiest to finish. One reason for this result was that this task had a shorter path than the paths for the *Vision* and *Vision with Sensory* tasks. The path distance for the *Vision* task was 840 inches (21 meters), for the
Vision with Sensory task 612 inches (15.3 meters) and 354 inches (8.85 meters) for the Sensor task. Another factor was the PDA processing time, which was fastest for the Sensor-only screen since this screen does not require the display of an image.

**Number of Errors and Precautions**

This section provides information regarding the number of errors and precautions during all three tasks. During all tasks and trials, there were no errors recorded. No software or hardware errors or failures occurred during the user study.

Table 13 provides the descriptive statistics across tasks and trials for the number of precautions issued during the trials. The term “precaution” as defined in Encarta World English Dictionary [33] is an action taken to protect against possible harm or trouble. In this situation, precaution is pressing the hardware stop button to protect the robot from crashing into walls or obstacles. The hardware stop button is pressed by a human near the robot. The raw data regarding the number of precautions for all participants is provided in Appendix F.

Table 13: The descriptive statistics related to the Number of Precautions for All Tasks across Trials

<table>
<thead>
<tr>
<th></th>
<th>Vision Task</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>2.20</td>
<td>2.40</td>
<td>3.23</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.88</td>
<td>2.21</td>
<td>1.91</td>
</tr>
<tr>
<td><strong>Trial 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>2.73</td>
<td>1.17</td>
<td>3.43</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>3.88</td>
<td>0.70</td>
<td>2.13</td>
</tr>
</tbody>
</table>
The average number of precautions for the *Vision* task during Trial one was 2.20. The average during Trial two was 2.73. Therefore, the number of precautions for the *Vision* task increased from Trial one to Trial two.

The average number of precautions for the *Sensor* task during Trial one was 2.40. During Trial two, this number decreased to 1.17. The number of precautions decreased dramatically between trials. This decrease occurred because the participants’ were permitted to directly view the environment and robot between trials.

The average number of precautions for the *Vision with Sensory* task during Trial one was 3.23. The average number of precautions during Trial two was 3.43. Therefore, the number of precautions for the *Vision with Sensory* task increased slightly between trials.

The number of precautions for all screens across trials is shown visually in Figure 37. According to the figure, the participants completed the task with the *Vision-only* screen with the least number of precautions, and recorded the highest number of precautions with the *Vision with Sensory* screen. During Trial two, the participants completed the task with the least number of precautions when using the *Sensor-only* screen.

Since the participants were permitted to view the robot and its environment during the *Sensor* task during Trial two, the number of precautions for the *Sensor* task during Trial two was the smallest compared to the other tasks.
Figure 37: Number of Precautions for all screens across trials

Number and Locations of Screen Touches

This section provides the analysis regarding the number and locations of screen touches for all three tasks.

Forward Button Screen Touches

Table 14 provides the descriptive statistics across tasks and trials for the number of screen touches for the move *Forward Button*. The detailed data is provided in Table G-1 of Appendix G.
Table 14: The descriptive statistics related to the Number of *Forward Button* Screen Touches for All Tasks across Trials

<table>
<thead>
<tr>
<th>Trial</th>
<th></th>
<th>Vision Task</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Average</td>
<td>8.33</td>
<td>5.73</td>
<td>6.00</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>3.30</td>
<td>1.96</td>
<td>2.36</td>
</tr>
<tr>
<td>2</td>
<td>Average</td>
<td>7.07</td>
<td>4.27</td>
<td>6.43</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>2.82</td>
<td>3.16</td>
<td>2.43</td>
</tr>
</tbody>
</table>

The average number of *Forward Button* screen touches for the *Vision* task during Trial one was 8.33. The average during Trial two was 7.07. Therefore, the number of *Forward Button* touches decreased from Trial one to Trial two.

The average number of *Forward Button* screen touches for the *Sensor* task during Trial one was 5.73. The average during Trial two was 4.27. The number decreased from Trial one to Trial two.

The average number of *Forward Button* screen touches for the *Vision with Sensory* task during Trial one was 6.00. The average during Trial two was 6.43. Therefore, the number increased between trials.

The number of *Forward Button* screen touches for all screens across trials is shown visually in Figure 38. According to the figure, the participants touched the *Forward Button* while using the *Vision-only* screen more than they did while using the other two screens, and they selected the *Forward Button* while using the *Sensor-only* screen less than they did while using the other two screens during Trial one. This number decreased even further during Trial two.
Figure 38: Number of *Forward Button* Screen Touches for all screens across trials

**Backward Button Screen Touches**

Table 15 provides the descriptive statistics across tasks and trials for the number of *Backward Button* screen touches. The detailed data is provided in Table H-2 of Appendix G.
Table 15: The descriptive statistics related to the Number of *Backward Button* Screen Touches for All Tasks across Trials

<table>
<thead>
<tr>
<th>Trial</th>
<th></th>
<th>Vision Task</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Average</td>
<td>0.37</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.67</td>
<td>1.19</td>
<td>0.85</td>
</tr>
<tr>
<td>2</td>
<td>Average</td>
<td>0.30</td>
<td>0.70</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>1.15</td>
<td>1.44</td>
<td>0.86</td>
</tr>
</tbody>
</table>

The average of *Backward Button* screen touches across all screens was less than 1. The number of selections during the *Vision* task during Trial one was 11 and there were 9 during Trial two. The number of *Backward Button* screen touches during the *Sensor* task for Trial one was 19 and for Trial two there were 21 selections. The number of *Backward Button* screen touches for the *Vision with Sensory* task during Trial one was 19 and 13 during the second trial. The numbers of *Backward Button* screen touches for all screens across trials are shown visually in Figure 39. According to the figure, there were slight decreases with the *Vision-only* and *Vision with Sensory Overlay* screens while there was a slight increase in the number of selections for the *Sensor-only* screen.
Figure 39: Number of *Backward Button* Screen Touches for all screens across trials

Right Button Screen Touches

Table 16 provides the descriptive statistics across tasks and trials for the number of screen touches for the *Right Button*. The detailed data is provided in Table G-3 of Appendix G.
Table 16: The descriptive statistics related to the Number of Right Button Screen Touches for All Tasks across Trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>Average</th>
<th>Vision Task</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>5.27</td>
<td>2.23</td>
<td>5.93</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>2.61</td>
<td>1.81</td>
<td>1.87</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>4.23</td>
<td>2.10</td>
<td>6.23</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>1.98</td>
<td>1.86</td>
<td>2.64</td>
</tr>
</tbody>
</table>

The average number of Right Button screen touches for the Vision task during Trial one was 5.27. The average during Trial two was 4.23. Therefore, the number of Right Button touches decreased from Trial one to Trial two.

The average number of Right Button screen touches for the Sensor task during Trial one was 2.23. The average during Trial two was 2.10. The number decreased from Trial one to Trial two.

The average number of Right Button screen touches for the Vision with Sensory task during Trial one was 5.93. The average during Trial two was 6.23. Therefore, the number increased between trials.

The numbers of Right Button screen touches for all screens across trials are shown visually in Figure 40. According to the figure, the participants touched the Right Button while using the Vision with Sensory Overlay screen more than they did while using the other two screens, and they selected the Right Button while using the Sensor-only screen less than they did while using the other two screens during Trial one. This result was also found during Trial two. The reason for touching the Right Button most while using the
Vision with Sensory Overlay screen could be the task description of this task. In task descriptions, the participants needed to turn the robot to the right.

![Graph showing number of screen touches across trials for Vision with Sensory Overlay, Vision-only, and Sensor-only screens.]

**Figure 40:** Number of Right Button Screen Touches for all screens across trials.

**Left Button Screen Touches**

Table 17 shows the descriptive statistics across tasks and trials for the number of screen touches for the Left Button. The detailed number data is provided in Table G-4 of Appendix G.
Table 17: The descriptive statistics related to the Number of Left Button Screen Touches for All Tasks across Trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>Average</th>
<th>Vision Task</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>6.43</td>
<td>4.53</td>
<td>4.67</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>2.62</td>
<td>2.83</td>
<td>2.58</td>
</tr>
<tr>
<td>2</td>
<td>Average</td>
<td>6.00</td>
<td>3.73</td>
<td>4.07</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>2.89</td>
<td>2.46</td>
<td>2.00</td>
</tr>
</tbody>
</table>

The average number of Left Button screen touches for the Vision task during Trial one was 6.43. The average during Trial two was 6.00. Therefore, the number of the Left Button touches decreased from Trial one to Trial two.

The average number of Left Button screen touches for the Sensor task during Trial one was 4.53. The average during Trial two was 3.73. The number decreased from Trial one to Trial two.

The average number of Left Button screen touches for the Vision with Sensory task during Trial one was 4.67. The average during Trial two was 4.07. Therefore, the number decreased between trials.

The number of Left Button screen touches for all screens across trials is shown visually in Figure 41. According to the figure, the participants touched the Left Button while using the Vision-only screen more than they did while using the other two screens, and they selected the Left Button while using the Sensor-only screen less than they did while using the other two screens during Trial one. This result is also found during Trial two.
Stop Button Screen Touches

Table 18 shows the descriptive statistics across tasks and trials for the number of screen touches for the Stop Button. The detailed data is provided in Table G-5 of Appendix G.
Table 18: The descriptive statistics related to the Number of *Stop Button* Screen Touches for All Tasks across Trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>Average</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vision</td>
<td>Sensor</td>
<td>Vision with Sensory</td>
</tr>
<tr>
<td></td>
<td>Task</td>
<td>Task</td>
<td>Task</td>
</tr>
<tr>
<td></td>
<td>26.10</td>
<td>15.03</td>
<td>25.27</td>
</tr>
<tr>
<td>2</td>
<td>Average</td>
<td>22.57</td>
<td>12.83</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>15.17</td>
<td>9.14</td>
<td>13.65</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>10.38</td>
<td>10.90</td>
<td>13.87</td>
</tr>
</tbody>
</table>

The average number of *Stop Button* screen touches for the *Vision* task during Trial one was 26.10. The average during Trial two was 22.57. Therefore, the number of *Stop Button* screen touches decreased from Trial one to Trial two.

The average number of *Stop Button* screen touches for the *Sensor* task during Trial one was 15.03. The average during Trial two was 12.83. The number decreased from Trial one to Trial two.

The average number of *Stop Button* screen touches for the *Vision with Sensory* task during Trial one was 25.27. The average during Trial two was 26.07. Therefore, the number increased between trials.

The number of *Stop Button* screen touches for all screens across trials is shown visually in Figure 42. According to the figure, the participants touched the *Stop Button* while using the *Vision-only* screen more than they did while using the other two screens, and they selected the *Stop Button* while using the *Sensor-only* screen less than they did while using the other two screens during Trial one. This result is changed during Trial two. The participants touched the *Stop Button* while using the *Vision with Sensory Overlay* screen more than they did while using the other two screens, and they selected
the Stop Button while using the Sensor-only screen less than they did while using the other two screens during Trial two.

![Graph showing the number of Stop Button screen touches for all screens across trials.](image)

Figure 42: Number of Stop Button Screen Touches for all screens across trials

No Button Screen Touches

Table 19 provides the descriptive statistics across tasks and trials for the number of screen touches for the No Button classification. The No Button classification includes all screen touches that did not select any of the buttons. The detailed data is provided in Table G-6 of Appendix G.
Table 19: The descriptive statistics related to the Number of No Button Screen Touches for All Tasks across Trials

<table>
<thead>
<tr>
<th></th>
<th>Vision Task</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.03</td>
<td>0.60</td>
<td>0.07</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.18</td>
<td>1.16</td>
<td>0.25</td>
</tr>
<tr>
<td>Trial 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.00</td>
<td>0.50</td>
<td>0.03</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.00</td>
<td>1.41</td>
<td>0.18</td>
</tr>
</tbody>
</table>

The average of No Button touches across all screens was less than 1. The number of touches during the Vision task during Trial one was 1 and there were no selections during Trial two. The number of No Button touches during the Sensor task for Trial one was 18 and for Trial two there were 15 touches. The number of No Button touches for the Vision with Sensory task during Trial one was 2 and only 1 during second trial. The numbers of No Button screen touches for all screens across trials are shown visually in Figure 43.

According to the figure, there were slight decreases with the Vision-only and Vision with Sensory Overlay screens while there was a decrease in the number of selections for the Sensor-only screen. An important result was that the participants touched the screen but not a button the most while using the Sensor-only screen. This result was found to hold true while using the other two screens during both trials. During both trials, the participants selected No Button while using the Vision-only screen fewer times than they did while using the other two screens.
Figure 43: Number of No Button Screen Touches for all screens across trials

Discussion

In general, the number of Stop Button touches was the highest, with an average of 21.31, followed by Forward Button touches, with an average of 6.31, followed by Left Button touches, with an average of 4.91, and finally followed by Right Button touches, with an average of 4.33. There were very few No Button touches, an average of 0.21, and Backward Button touches, with an average of 0.51. The reason for more Forward Button touches than turning and Backward Button touches is that all tasks include going forward more than turning right or left, and none of the tasks required the backward motion. The reason for more Left Button touches than Right Button touches is that there were more left turns than the right turns in the user study tasks.
**Accuracy of Goal Achievement**

This section provides information regarding the accuracy of the participants to achieve the defined goal locations for all three tasks. Table 20 shows the descriptive statistics across tasks and trials. The detailed accuracy data for all participants is provided in Appendix H.

The distance to the goal point was measured after each task and if the vertical and/or horizontal distance to the goal point was less than or equal to 24 inches and its horizontal distance was more than 12 inches, then the position was considered as *almost reached* the goal point. Figure 44 shows the illustration of the vertical and horizontal distances. If the vertical distance to the goal point was 0 inches and the horizontal distance was less than or equal to 12 inches, then the robot *reached* the goal point. If the robot’s front side passed the goal point then the robot *passed* the goal location. When this occurred, the robot’s body was usually on top of the goal point. Otherwise, the robot’s goal achievement accuracy was considered to be *not reached*.

![Figure 44: Illustration of Vertical and Horizontal Distance](image-url)
Table 20: The mean of the Goal Achievement Accuracy for All Tasks across Trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>Vision Task</th>
<th>Sensor Task</th>
<th>Vision with Sensory Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reached</td>
<td>Not Reached</td>
<td>Almost Reached</td>
</tr>
<tr>
<td>Trial 1</td>
<td>12</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td>Trial 2</td>
<td>23</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>18</td>
<td>2</td>
</tr>
</tbody>
</table>

The total number of the participants that reached or almost reached the goal location for all tasks during both trials is shown visually in Figure 45. As can be seen from the figure, the number of participants that reached or almost reached the goal location during the Vision with Sensory Overlay task is very low. The reason for this result is the long processing time to display information on this screen which caused communication and commanding delays. Since Vision with Sensory Overlay screen combines the visual, laser, and sonar information into a single screen, the update time is quite long for the screen. Therefore, there is a delay between the participants’ issuing commands and robot’s action.
Figure 45: Accuracy of the Goal Achievement for the *Reached* and *Almost Reached* for all tasks during both trials

**Vision Task**

The accuracy of the goal achievement for the *Vision* task during Trial one can be summarized as follows: Twelve participants were able to *reach* the goal point. Seven did *not reach* the goal point. Nine participants *almost reached* the goal point while two participants *passed* the goal point.

The goal achievement accuracy for the *Vision* task during Trial two is summarized as follows: Twenty-three participants *reached* the goal point. Two did *not reach* the goal point. Four participants *almost reached* the goal point. Only one participant *passed* the goal point.
The difference between the goal achievement accuracy between trials for the Vision task is shown visually in Figure 46. The number of participants that reached the goal location condition increased dramatically during Trial two. The reason for this increase may be related to the learning effects and more experience with the system.

![Figure 46: Accuracy of the Goal Achievement for the Vision task during both trials](image)

Sensor Task

The accuracy of the goal achievement for the Sensor task during Trial one can be summarized as follows: Fourteen participants were able to reach the goal point. Eight did not reach the goal point. Six participants almost reached the goal point while two participants passed the goal point.
The goal achievement accuracy for the *Sensor* task during Trial two is summarized as follows: Twenty-six participants *reached* the goal point. One participant did *not reach* the goal point. Only one participant *almost reached* the goal point. Two participants *passed* the goal point.

The difference between the goal achievement accuracy between trials for the *Sensor* task is shown visually in Figure 47. The number of participants that *reached* the goal location condition increased dramatically during Trial two. The reason for this increase was that the participants were permitted to view the robot and its environment during Trial two.

![Figure 47: Accuracy of the Goal Achievement for the Sensor task during both trials](image-url)
Vision with Sensory Task

The accuracy of the goal achievement for the Vision with Sensory task during Trial one can be summarized as follows: Three participants were able to reach the goal point. Twenty-three did not reach the goal point. Four participants almost reached the goal point while none of the participants passed the goal point.

The goal achievement accuracy for the Vision with Sensory task during Trial two is summarized as follows: Seven participants reached the goal point. Eighteen did not reach the goal point. Two participants almost reached the goal point while three participants passed the goal point.

The difference between the goal achievement accuracy between trials for the Vision with Sensory task is shown visually in Figure 48. The number of participants that reached the goal location increased during Trial two. A large number of people could not reach the goal location. The reason for this result was the long processing time in the Vision with Sensory Overlay screen. This screen provides the camera image, the laser data, and the sonar data in a single screen; therefore it takes more time to update the vision and sensory data.

As it can be seen from the Figure 48, the number of participants that reached the goal location during the Vision with Sensory Overlay task is very low. The long processing time is the reason for this result. While participants were using the Vision with Sensory Overlay screen, there existed a long delay between the participants’ commands and robot’s action which affected this result.
Figure 48: Accuracy of the Goal Achievement for the Vision with Sensory task during both trials

The Reached Condition across All Tasks

The accuracy of the goal achievement difference for the reached condition across all tasks during both trials is shown visually in Figure 49.
During Trial one, the goal achievement accuracy for the *reached* condition during the *Sensor* task was the highest when compared to the other tasks. The *Vision with Sensory* task had the lowest accuracy of the three tasks. During Trial two, the accuracy for the *reached* condition was again the highest for the *Sensor* task and the *Vision with Sensory* task was the lowest.

The *Not Reached* Condition across All Tasks

The accuracy of the goal achievement difference for the *not reached* condition across all tasks during both trials is shown visually in Figure 50.
During Trial one, the goal achievement accuracy for the *not reached* condition during the *Vision with Sensory* task was the highest when compared to the other tasks. The *Vision* task had the lowest accuracy of the three tasks. During Trial two, the accuracy for *not reached* condition was again the highest for the *Vision with Sensory* task and the *Sensor* task was the lowest.

**The Almost Reached Condition across All Tasks**

The accuracy of the goal achievement difference for the *almost reached* condition across all tasks during both trials is shown visually in Figure 51.
During Trial one, the goal achievement accuracy for the *almost reached* condition during the *Vision* task was the highest when compared to the other tasks. The *Vision with Sensory* task had the lowest accuracy of the three tasks. During Trial two, the accuracy for the *almost reached* condition was again highest for the *Vision* task and the *Sensor* task was the lowest.

The *Passed* Condition across All Tasks

The accuracy of the goal achievement difference for the *passed* condition across all tasks during both trials is shown visually in Figure 52.
During Trial one, the goal achievement accuracy for the passed condition during the Vision and Sensor tasks were the higher than the Vision with Sensory task. During Trial two, the accuracy for the passed condition was the highest for the Vision and Sensor task and the Vision task was the lowest.

Discussion

In general, the goal achievement accuracy for the reached condition during the Sensor task was the highest and the Vision with Sensory task had the lowest accuracy of the three tasks. One reason for this result was that the Sensor task had a shorter path than the paths for the Vision and Vision with Sensory tasks. The path distances were provided...
Another factor was the PDA processing time, which was fastest for the Sensor-only screen since this screen does not require the display of an image. Therefore, participants were able to successfully finish this task over the other two tasks. Another factor was that there was a human in the hall telling the participant to stop during the Sensor task.

The participants performed poorly with the Vision with Sensory Overlay screen. The reason for this result was the long processing time that delayed communications between the PDA and the robot. This delay between the participants’ commands and robot’s action requires more time to drive the robot to the goal location.
CHAPTER VII

CONCLUSIONS AND FUTURE WORK

In this thesis, a PDA-based teleoperation interface for a mobile robot and the results of a user evaluation with thirty novice users were presented. In this interface, three touch-based PDA screens were developed: a Vision-only screen, a Sensor-only screen, and a Vision with Sensory Overlay screen. The following hypotheses were defined for the user evaluation. The first was: “The Vision with Sensory Overlay screen and Sensor-only screen induce higher workload levels for the defined user group than the Vision-only screen.” The second was: “The Vision-only screen is easier to use for the defined user group than the Vision with Sensory Overlay screen and Sensor-only screen.” The third was: “The Vision-only screen is the preferred teleoperation interface by the defined user group.” The fourth was: “The Vision with Sensory Overlay screen is preferred for teleoperation over the Sensor-only screen by the defined user group.” The results of the user evaluation were statistically analyzed.

Conclusions

Workload Hypothesis Analysis

The first hypothesis relies on the workload analysis results. The first hypothesis claimed that: “The Vision with Sensory Overlay screen and Sensor-only screen induce higher workload levels for the defined user group than the Vision-only screen.” When participants are required to use all three screens in a remote location, the Vision-only
screen requires the least amount of workload. This result verifies the first hypothesis for this condition. The hypothesis also holds for the Vision-only and Vision with Sensory Overlay screens during the second trial.

The results indicating that the perceived workload was significantly lower for the Sensor task than the other two tasks when the participants could directly view the environment and robot are not surprising. In fact, this result would be expected. What cannot be determined by this study is if the above stated hypothesis would be valid when the participants were permitted to directly view the environment for all three tasks. It is anticipated that the results would be similar, but with lower workload results, to those found during the first trial of this evaluation.

Ease of Use Hypothesis Analysis

The second hypothesis relies on the Likert scale and Ranking questions analysis results. Hypothesis two states: “The Image-only screen is easier to use for the defined user group than the Image with Sensory Overlay screen and Sensor-only screen.” When the participants are required to use all three screens from a remote location, they rated the Vision-only screen as significantly easier to use than the Sensor-only and the Vision with Sensory Overlay screens in the Likert scale questions part. Participants also found it significantly easier to correct their errors with the Vision-only screen than the Sensor-only screen and the Vision with Sensory Overlay screen when all tasks were executed from a remote environment. As the result of the Ranking questions analysis, the participants ranked the Vision-only screen as significantly easier to use than the Sensor-only and the Vision with Sensory Overlay screens when they are required to use all three screens from
a remote environment. These results provide verification of hypothesis two for this execution condition of the tasks.

For both the Likert scale and Ranking questions analysis, when participants’ environmental condition changed during Trial two of the Sensor task, the verification of the hypothesis no longer holds. This would be expected, since during Trial two of the Sensor-only task, the participants were permitted to directly view the robot and its environment, therefore the Sensor-only screen was ranked as easiest to use. During Trial two it was found that Vision-only screen was significantly easier to use than the Vision with Sensory Overlay screen, therefore the portion of the hypothesis related to these two screens still holds.

Overall Preferred Screen Hypothesis Analysis

The third hypothesis relies on the Ranking questions analysis results. The hypothesis was: “The Image-only screen is the preferred teleoperation interface by the defined user group.” The participants gave a significantly higher general overall ranking to the Vision-only screen than the Sensor-only and Vision with Sensory Overlay screens when the participants were required to use all three screens from a remote location. This provides the verification of the third hypothesis. During the second trial, the results do not support the third hypothesis.

Sub-Screen Preference Hypothesis Analysis

The fourth hypothesis relies on the Ranking question analysis results. Unfortunately, this hypothesis could not be completely verified through this evaluation. It was claimed that “The Image with Sensory Overlay screen is preferred for teleoperation
over the *Sensor-only* screen by the defined user group.” The results show that during Trial one, participants provided overall rankings of 1.63 for the *Sensor-only* screen and 1.87 for the *Vision with Sensory Overlay* screen. The results reversed for the *Sensor-only* and the *Vision with Sensory Overlay* screens during Trial two. The participants provided overall rankings of 1.93 for the *Sensor-only* screen and 1.83 for the *Vision with Sensory Overlay* screen. Neither of these two results was significant. This difference in the data during the second trial would be expected, since during this trial the participants were able to see the robot and environment while completing the *Sensor* task. As an overall result, the fourth hypothesis is not verified.

**Future Work**

The following future work can be added to the work presented in this thesis.

- The ability to pan and tilt the camera can be added to the vision based screens: *Vision-only* and *Vision with sensory Overlay* screens. This may make these screens more useful.

- Another camera can be mounted on the back of the robot.

- The robot icon in the *Sensor-only* screen could be properly scaled so that distances on the screen are more intuitive.

- The laser and sonar data points could remain displayed until the new points are ready to be displayed on the *Sensor-only* and *Vision with Sensory Overlay* screens.

- The ability to turn the robot a pre-determined number of degrees right or left could be added to the interface. That way, regardless of how fast the screen updates, the robot would turn a certain amount and stop on its own.
• Potential options for reducing the delays and speeding up the processing can be investigated. One option may be reprogramming the processes in another language as well as optimizing the code. Another option may be using a PDA that has more powerful processor.