

Novel Egocentric Robot Teleoperation Interfaces for Search and Rescue

by

Stela Hanbyeol Seo

A thesis submitted to
Faculty of Graduate Studies of
University of Manitoba
in partial fulfillment of the requirements
of the degree of

Doctor of Philosophy

Department of Computer Science
University of Manitoba
Winnipeg, Manitoba, Canada

© Copyright 2021 by Stela H. Seo

Novel Egocentric Robot Teleoperation Interfaces for Search and Rescue

Abstract

Teleoperation is a powerful tool: a person can have conferences and meetings overseas, visit their families abroad, go to uncharted locations, and explore dangerous environments. However, during real-time remote teleoperation, the operator faces various challenges every moment, primarily maintaining remote awareness and performing under high cognitive load. The challenges in search and rescue teleoperation exacerbate its difficulty.

To successfully teleoperate a remote robot and accomplish tasks, the operator must maintain a high level of situation awareness by understanding the remote robot's current states and the environment while being aware of their mission tasks. However, the operator has a limited access to the remote environment through teleoperation interfaces. Additionally, the interfaces can only deliver limited data (i.e., limited field of view and limited types of sensors). To make things worse, in search and rescue teleoperation, the operator must make important decisions that may impact victims' life with the limited information.

As remote teleoperation interfaces are the only gateway for most of the time, how they deliver the information impacts the operator's situation awareness. Researchers found that the ways of presenting information impact users' overall task performance in terms of accuracy, completion time, and workload in human-computer interaction and human-robot interaction. We extend this theme to search and rescue teleoperation scenarios.

We explore novel interface designs to support the operator by retrieving remote information and presenting them in a way that the operator can understand in time. Our design helps the operator increase their situation awareness and their overall task performance. We further discuss benefits and drawbacks of our implementations further contributes to improving future teleoperation interface designs.

Contents

Abstract	i
Contents	iii
List of Figures	ix
List of Tables	xix
Publications	xxi
Acknowledgements.....	xxiii
Chapter 1 Introduction	1
1.1. Teleoperation Challenges.....	4
1.2. Investigating Novel Interface Strategies to Improve Teleoperation Performance and Experience.....	7
1.2.1. Improving Teleoperation Effectiveness with Multiple Cameras	10
1.2.2. Providing the Location Information of Team Robot Members	12
1.2.3. Exploring Interface Design Techniques for Graphical Representations of Team Robot States	15

1.2.4. Informative Visual Transition in Multi-robot Teleoperation.....	18
1.2.5. Summary	20
1.3. Contributions.....	21
Chapter 2 Background and Related Work	23
2.1. Situation Awareness, Teleoperation Challenge	25
2.2. From Human-computer Interaction and Human-robot Interaction to Novel Teleoperation Interfaces.....	29
2.2.1. Improving Teleoperation Interface Designs	32
2.2.2. Multi-video Stream in Egocentric Video Teleoperation Interfaces	35
2.2.3. Group Awareness in Multi-robot Teleoperation.....	36
2.3. Summary	38
Chapter 3 Multiple Cameras: Interactive Detail-in-context Interface Using Multiple Cameras in Teleoperation.....	41
3.1. Monocle, the Detail-in-context Interactive Interface.....	46
3.2. Implementation	47
3.3. Evaluation: Formal User Study.....	49
3.3.1. Comparison Interfaces: Toggle and Side-by-side Views.....	50
3.3.2. Task.....	51
3.3.3. Measures	52
3.3.4. Procedure	54

3.3.5. Apparatus	55
3.4. Results.....	56
3.5. Discussion	60
3.6. Future Work	61
3.7. Summary	62
Chapter 4 Team Location: In-feed Embedded Techniques for Visualizing Robot Team Member Locations.....	65
4.1. Designing In-feed Location Indicators	68
4.2. Our In-feed Team Robot Location Indicators.....	70
4.3. Teleoperation Adaption Challenges.....	71
4.4. Implementation	73
4.5. Evaluation: The Formal User Study	73
4.5.1. Apparatus	74
4.5.2. Task.....	75
4.5.3. Measures	77
4.5.4. Procedure	78
4.6. Results.....	79
4.6.1. Quantitative Results	79
4.6.2. Qualitative Results	81
4.7. Discussion	83

4.8. Summary	84
Chapter 5 Team States: Graphical Techniques for Maintaining Awareness of Team Robots.....	87
5.1. Graphical Representations of Robotic States in Multi-Robot Team Teleoperation	89
5.1.1. Template Generic Team Robot States	90
5.1.2. On-Screen Visual Representation Strategy.....	92
5.1.3. Graphical Visualization Parameters.....	95
5.2. Exploration Testbed	97
5.2.1. Robot Team-member Simulation.....	98
5.2.2. Navigation Task.....	99
5.2.3. Action Task.....	100
5.2.4. Measurements	100
5.2.5. Implementation	101
5.3. Design Exploration	102
5.3.1. Exploration 1.....	103
5.3.2. Exploration 2, Two Team Robots.....	109
5.3.3. Exploration 3, Greyscale variants.....	111
5.3.4. Exploration 4, Improved State Questionnaire.....	112
5.3.5. Exploration 5, Alternative Questionnaire	115

vii

7.2. Envisioning the Teleoperation Consequences	162
7.3. Moving Forward from Our Work	163
7.4. Our Findings and Contributions.....	165
7.5. Toward Advanced Teleoperation Interfaces	166
Bibliography	168
Appendix A Multiple Cameras: Interactive Detail-in-context Interface Using Multiple Cameras in Teleoperation.....	177
Appendix B Team Location: In-feed Embedded Techniques for Visualizing Robot Team Member Locations.....	179
Appendix C Team States: Graphical Techniques for Maintaining Awareness of Team Robots.....	181
Appendix D Control Transition: Utilization of Visual Transition as an Information Source in Multi-robot Teleoperation	183

List of Figures

Figure 1-1. An example egocentric video teleoperation interface with an on-head camera (left). The video feed does not include the robot (right, mock-up), sending the video feed.....	3
Figure 1-2. This research covers the breadth of the teleoperation challenges with the perspective of numbers of robots (horizontal) and information varieties (vertical). The alphabet letters are the order of our introduction in this document.	8
Figure 1-3. Canadian Army’s unmanned ground vehicle (left) and controller box (right). The controller box contains camera buttons (yellow ellipse) on top of the robot diagram to switch camera feed. The monitor on the controller box displays one video feed at a time (i.e., toggle view).....	11
Figure 1-4. Versatrax 450 Robot Crawler (left) from Inuktun Inc. and its control box (right). The control box can display two camera feeds from the robot at once (side-by-side view).	12
Figure 1-5. Our novel teleoperation interface (bottom) using two live camera feeds from a single robot. The interface merges two camera views of a single robot (top) into one mixed-resolution, detail-in-context view (bottom).	13

Figure 1-6. A mini-map example without the environmental data. It is on the remote video feed's right. There are four robots in the scene (black, red, blue, green ellipses). 14	
Figure 1-7. One of our in-feed location indicators. It takes significantly less screen space compared to a mini-map. The implementation detail is in Chapter 4..... 15	
Figure 1-8. An operator teleoperates a robot located in remote environments, with two additional robot team members. While controlling one robot, they maintain awareness of other robots using on-screen widgets placed at the bottom of the screen. 17	
Figure 1-9. Examples from our exploration space. We have three information representation techniques (text, icon, emotional information encoding) and graphical visualization parameters (color coding, animation, and the number of team robots). The detail can be found in Chapter 5..... 18	
Figure 1-10. An operator remotely controls a robot to avoid debris (left); meanwhile, another robot (right) requests an operator's attention to work on a complex object. To help the robot on the right, first the operator must understand the situation around the robot. 19	
Figure 1-11. A video game's loading screen with additional background information. Dragon Age: Inquisition, Electronic Arts, 2014. 20	
Figure 2-1. Model of Situation Awareness in dynamic decision making (Endsley, 1995). The figure is from the publication. 27	
Figure 2-2. Two examples of improving visualization on teleoperation interfaces. The left interface integrates two cameras' feeds to show more information to the operator (Seo, Rea, et al., 2017). The right interface augments extra visual objects to help	

the operator locate the critical information quickly (Rea et al., 2017).	29
Figure 2-3. An example of video centric teleoperation interfaces (Singh et al., 2013). The interface primarily has a remote video feed on the screen and has other information around the video feed.....	34
Figure 3-1. The mobile robot in the remote space (right) constantly provides sensor inputs to the teleoperation system (left) in real-time. With a toggle view interface, only one view is visible and other views are hidden. This figure conceptualizes the toggle view by putting one hidden view at other's back on the teleoperator's monitor..	43
Figure 3-2. The mobile robot in the remote space (right) constantly provides sensor inputs to the teleoperation system (left) in real-time. With a tile (or side-by-side) view interface, on the teleoperator's monitor in this figure, all views are visible with reduced resolution on the screen.....	43
Figure 3-3. Monocle, our detail-in-context interface, a fully interactive detail-in-context teleoperation interface with wide pannable (behind the robot) and narrow pan-and-tilt (in the head) views. An operator can move the monocle around to gain detail where needed. (Low light, image quality due to search and rescue scenario).....	44
Figure 3-4. Monocle detail-in-context interface. (left) the sharp view integrates into the wide view; (right) the monocle can be toggled off by the operator. The third-person view provides a broad context, while the first-person view has a narrow field-of-view with more clarity.	45
Figure 3-5. Raspberry Pi's camera with no additional filter (left) and 180° fish-eye lens (right). With a fish-eye lens, the robot can capture more remote information; however, the details are missing since the camera's sensor and resolution are the	

same.	48
Figure 3-6. The robot with the boom; the boom camera is above the robot and points downward to capture the full body. Both cameras have LED lights attached to aid in dark environments.....	49
Figure 3-7. Sharp narrow view (left), wide-angle view (right). The green arrow indicates the robot’s facing direction (in contrast to viewing direction), blue border at the edge of the screen appears to indicate camera movement limitations. Both cameras’ feeds are shown together in the side-by-side (tile) interface, but individually for the toggle interface. Note the victim, with simulated red blood, is easier to see in the sharp view.	51
Figure 3-8. The picture (left) and map (right) of our search and rescue evaluation environment. The robot is at the start position in the picture. On the map, blue boxes are equipment, yellows are rubbles, the black surfaces are (fake) holes to avoid stepping on, and the purple box is a wall.....	52
Figure 3-9. A participant sits at the desk (left) and teleoperates a mobile robot located in a different room using a PC and a dual-axis Xbox 360 controller (middle). We printed out the control scheme (right) and placed it on the desk for participants to use during the study.....	56
Figure 3-10. Overall average performance scores per interface with 95% confidence interval. The data skewed to negative values as there are many positive outcomes while infinite negative outcomes in our performance score aggregation.	57
Figure 3-11. The graph illustrates how many participants did best, second best, or worst with each interface. For example, one participant performed worst with Monocle,	

while six participants performed worst with toggle.....	58
Figure 3-12. Average performance scores with breakdown of positive and negative factors.	
The number of accidents was the largest difference between interfaces.....	61
Figure 4-1. An example illustration when a team robot (right) finds a blockage and calls for help. If an operator, who controls a nearby (left) robot, is aware of the team member's location, they can quickly start moving toward the team member and save time to clear the path.....	
	67
Figure 4-2. An example of mini-map on an Android phone. The mini-map takes a portion of the screen. Image retrieved from forum.xda-developers.com on August 7, 2020.	
.....	67
Figure 4-3. Two examples of off-screen target markers in video games. In the left example, off-screen targets (other players) are highlighted in red (<i>Overwatch</i> , <i>Blizzard</i> , 2016).	
The right example shows angled pointy markers at the edge of the mini-map with different colors for different types of off-screen targets (<i>Final Fantasy XIV</i> , <i>Square Enix</i> , 2013).....	
	69
Figure 4-4. (left) Wedge (Gustafson et al., 2008) in mobile human-computer interaction literature shows off-screen objects. (right) We adapted and built an in-feed embedded indicator in first-person teleoperation view, showing team robot members' locations.	
	71
Figure 4-5. Our in-feed team robot location indicators on egocentric teleoperation interfaces. A and B show mock-ups of how Halo and Wedge work; only the viewport, with the indicators at the edges, are shown. C shows arrows that indicate closer objects with longer tails. In this scene, the blue robot is right and slightly in	

front of the robot, teleoperated by an operator.	72
Figure 4-6. Interface used by participants. Participants perform the distractor task (searching for green lights, one shown, indicated by a yellow arrow in this figure) and hit the spacebar when one was found (score shown on the right), while maintaining awareness of the other robot members' location information. The mini-map was shown only in the mini-map case, with the region left black in the other cases.	75
Figure 4-7. Screenshot of the Wedge condition interface used in the user study. In non-mini-map conditions, the mini-map (shown in Figure 4-6) was removed and the region filled black.	77
Figure 4-8. Charts represent relative performance showing how many participants performed, best or worst, with a specific interface.	80
Figure 4-9. All scales were 0~20, with higher scores indicating more attention demanded and higher confidence. Confidence was measured twice (once for closest, once for furthest, robot) and summed. Error bars indicate 95% confidence intervals.	81
Figure 5-1. An operator teleoperates a robot located in remote environments, with two additional robot team members. While controlling one robot, they maintain awareness of other robots using on-screen widgets (screen bottom).	89
Figure 5-2. A team robot indicator with its moving direction (arrow on the bottom) and gaze direction (cone on the top).	92
Figure 5-3. Text representations for all states conveyed. Each state consists of two words: a noun and a descriptive adjective. Text is tilted to maximize size within a square space, the same size with icon representation (see Figure 5-4). We used a black	

outline to increase contrast and maximize visibility.....	93
Figure 5-4. Icon representations have one-to-one mappings to the text (see Figure 5-3). The icons are placed around the robot model, in the same position as the text indicators.	94
Figure 5-5. Emoji shows the condition of general states using a facial expression. The message state is embedded in facial expressions. The bottom example shows that a team robot has all positive states and no messages.....	95
Figure 5-6. Every asset in the greyscale version: text, icon, and emoji representations. In this version, the team robot is in greyscale. Examples with each representation are on the right.	96
Figure 5-7. Two sprites (happy with an urgent message and angry with no messages) – Each sprite has 40 frames, which provide smooth animation. One cycle of the animation takes about 1.33 seconds. The animation continually plays in our animated representations.....	97
Figure 5-8. Our team state representations for two team robots with emotional information encoding. Note that the message state is embedded onto the face.....	98
Figure 5-9. A participant is unloading the robot. The mission goal and its progression are shown at the top center of the screen. Note that while the goal is always displayed, the progress bar only appears when the robot is at the designated location.	99
Figure 5-10. A top-down map of the environment. The map shows two underpasses (desks, lighter shading), obstacles, and designated locations. The map is always on the secondary monitor during studies, but the robot’s current location is not marked.	100

Figure 5-11. Our on-screen questionnaire to measure a participant’s team state awareness.

We did not include the color coding in the questionnaire. When the questionnaire pops up, the teleoperation interface blacks out its screen, stops all processes, and waits for the completed questionnaire. 101

Figure 5-12. Jackal UGV robot from Clearpath Robotics Inc. (left) and ThrustMaster HOTAS Warthog Flight Stick (right) used in our design explorations. 102

Figure 5-13. Exploration 1 NASA TLX performance subscale report. Error bar indicates the standard error. Lower scores are more positive. That is, the participant thought that they were more successful. 108

Figure 5-14. Our on-screen questionnaire to measure a participant’s team state awareness for two team robots. See Appendix C for the high-resolution captures..... 109

Figure 5-15. Our modified questionnaire with the icons that match the widget designs. We show a portion of the questionnaire for icons only (A) and emojis (B) for convenience. See Appendix C for the high-resolution captures. 113

Figure 5-16. Exploration 4 NASA TLX effort subscale report. Error bar indicates the standard error. Lower scores are more positive. 115

Figure 5-17. Our stripped-down questionnaire, which only asks one state of one team robot and the robot’s general well-being..... 116

Figure 6-1. The concept of the control transition in multi-robot teleoperation. At the current point on the time domain (green vertical har), the operator switches their control from one robot to another. That is, robot A is no longer under a control of the operator after this point. 125

Figure 6-2. Example loading screens from games. Left loading screen provides a tip of

moving from one space to another using the door animation (<i>Resident Evil</i> , <i>Capcom</i> , 1996). The right loading screen gives a gamer the next stage's context information (<i>Medal of Honor: Allied Assault</i> , <i>Electronic Arts</i> , 2002).	126
Figure 6-3. The concept of the control transition. Green vertical bar conceptualizes the current point in time when the operator must switch their control from robot A to robot B.	130
Figure 6-4. A part of a map in a video game with the player's trail. The trail starts at the highlighting circle on the top, passes the sword icon (a battle scene), and continue stretching downward. Screenshot from <i>Dragon Age: Origins</i> , <i>BioWare</i> , 2009.	142
Figure 6-5. Our initial design framework. The blue box contains the type of information, the green contains teleoperation information, and the yellow contains groups of techniques taken from other fields.	145
Figure 6-6. Multi-UAV testbed interface (Calhoun et al., 2012; Draper et al., 2008)....	149
Figure 6-7. RoboLeader multi-robot teleoperation interface (J. Y. C. Chen & Barnes, 2012a, 2012b; J. Y. C. Chen et al., 2010).	151
Figure 6-8. A multi-robot teleoperation interface by SMP Robotics Systems Corp. Retrieved image from smprobotics.com/products_autonomous_ugv/multi-robot-control-system/ on April 20, 2020.	153
Figure 7-1. Examples of our design implementations from our projects.	166

List of Tables

Table 3-1. Summary of trade-offs emerging from the participant feedback analysis.	61
Table 5-1. General robot states. We explore variations of the four robot states: connectivity, battery level, physical damage, and mission messages. We abstract the details into three categories.	91

Publications

Some ideas and figures in this thesis have appeared previously in the following publications by the author.

STELA H. SEO, Daniel J. Rea, Joel Wiebe, and James E. Young. 2017. “Monocle: Interactive detail-in-context using two pan-and-tilt cameras to improve teleoperation effectiveness.” In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 962–967. doi.org/10.1109/RO-MAN.2017.8172419

STELA H. SEO, James E. Young, and Pourang Irani. 2017. “Where are the robots? In-feed embedded techniques for visualizing robot team member locations.” In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 522–527. doi.org/10.1109/ROMAN.2017.8172352

STELA H. SEO, James E. Young, and Pourang Irani. 2020. “How are Your Robot Friends Doing? A Design Exploration of Graphical Techniques Supporting Awareness of Robot Team Members in Teleoperation.” *International Journal of Social Robotics* (2020). doi.org/10.1007/s12369-020-00670-9

Acknowledgements

I would like to express my deepest gratitude to all people supported me throughout my PhD progress. Foremost, I would like to explicitly thank my family – Michael, Gabriella, Yuri, Han, Rosa, Grace, and Agape. Without their consideration, assistance, and love, I would not be able to make it this far.

I would like to thank Dr. James E. Young, the thesis advisor; Dr. Pourang Irani and Dr. Jason Morrison, the internal PhD advisory committee members; and Dr. Derek Reilly, the external PhD examiner for their consideration, support, and guidance. Their guidance helps me develop my PhD thesis as well as advance as a researcher.

The department of Computer Science at the University of Manitoba has been my home for many years. Faculty members always have welcomed me for having a discussion in various topics. Support staffs have always provided me invaluable assistance throughout my academic programs. Lab mates in the human-computer interaction laboratory and the human-robot interaction laboratory have always helped me improve as a researcher.

Community members from St. Paul the Apostle Parish have always encouraged me and given me the positive energy. My friends, fellow researchers, and relatives overseas have cheered me up for the successful PhD dissertation. It would be an endless list if I start

listing all people in my mind; I really appreciate their assistance and encouragement.

Lastly, I would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and the University of Manitoba for their financial assistance. Their financial supports were essential during my PhD program.

Chapter 1 Introduction

Search and rescue robots are no longer science fiction. We can buy an affordable search and rescue robot from many robotics companies. We have seen working robots in disaster sites such as the Notre Dame Blaze (Pescoe-Yang, 2019), the Fukushima Nuclear Disaster (BBCClick, 2017), and the World Trade Centre Tragedy (Casper & Murphy, 2003), where the crews teleoperate robots to get to dangerous areas from safe and distant locations.

Remotely controlling a robot is a challenging task. Teleoperators need to be aware of their robots to make decisions impacting their missions in real-time (Phillips & Jentsch, 2017). In search and rescue teleoperation, the difficulty escalates as operators must simultaneously perform multiple tasks. These include avoiding debris in a dangerous environment, finding victims, inspecting equipment, maintaining team communication, and controlling a group of robots while pursuing mission objectives (Demir, McNeese, & Cooke, 2017; Lewis, 2013; Rea, Seo, Bruce, & Young, 2017; Seo, Rea, Wiebe, & Young, 2017).

For search and rescue, having visual access to the remote environment is critical to the task (Rea et al., 2017). However, as the environment may be dangerous for people to physically get to (Casper & Murphy, 2003), it is safe to gain visual access by cameras remotely. Using installations (e.g., surveillance video cameras) for visual access may not be reliable because of unexpected changes to the environment caused by disasters, collapses, or hazards (Liu & Nejat, 2013). These conditions have led to the current and predominant control paradigm for teleoperation: the operator sits at a desk, watches video feeds from on-robot cameras, and provides commands to the robot using an egocentric video teleoperation interface (Casper & Murphy, 2003; Liu & Nejat, 2013; Nielsen, Goodrich, & Ricks, 2007; Singh et al., 2013).

Egocentric video teleoperation interfaces (e.g., Figure 1-1) allow an operator to



Figure 1-1. An example egocentric video teleoperation interface with an on-head camera (left). The video feed does not include the robot (right, mock-up), sending the video feed.

control the robot as if they are in or with the robot. For example, when an operator drives the remote robot forward, all sensors, including video cameras, move forward together with the robot because they are on the robot. Hence, all information that the operator receives through the interface also changes accordingly.

To support the operator's understanding of the remote environment, the mobile robots constantly provide sensor readings to the operator. For example, a remote video stream (Casper & Murphy, 2003), possibly the robot within that environment (Saakes et al., 2013), helps the operator understand the remote environment and the robot's status during teleoperation. However, people have limited capacity to comprehend information simultaneously (Endsley, 2016). We need a good strategy to present the constant streams of data to the operator so that they can comprehend the information in real-time teleoperation.

Improving interface designs is a successful strategy for supporting teleoperators' understanding of the remote situation and increasing their task performance. Well-designed interfaces can help operators reduce their cognitive load and improve their overall task

performance (Singh et al., 2013). This thesis explores novel egocentric teleoperation interface designs to help the operator's awareness and their tasks in search and rescue scenarios.

1.1. Teleoperation Challenges

Teleoperation is to maneuver a device or machine with electric signals at a distance. The idea was introduced in the 1800s (613809, 1898). Application domains expand to business, education, medical science, film, the military, and so on, where it includes remote operations at various distances, for example, controlling a collocated robot (Wentzel, Rea, Young, & Sharlin, 2015) or a rover on Mars (Bajracharya, Maimone, & Helmick, 2008). In this thesis, teleoperation controls one or more robots from a remote location where the operator cannot see the robots (e.g., behind walls, across office rooms, or overseas).

Controlling a robot in a separated space is challenging: an operator must understand the remote environment using the data from the robot solely, because the operator cannot use their biological senses (e.g., vision, sound, touch, smell, heat, balance, etc.) in teleoperation unlike being in person in the place. For example, many people can differentiate floorings (e.g., carpet and hardwood) while walking on the floor without visual inspection. An operator must visually analyze or comprehend sensor data to understand even the subconsciously understandable information in teleoperation.

For successful teleoperation, an operator needs to maintain a high level of situation awareness (Phillips & Jentsch, 2017; Seo, Rea, et al., 2017), which helps the operator improve their overall performance (Endsley, 2016). Situation awareness explains the operator's understandability of the remote environment (Endsley, 2016). This awareness also includes their understandability of mission goals, internal and external states of their robots,

physical configurations, events or changes in the remote environment, and future expectations (Endsley, 2016; Yanco & Drury, 2004b).

One way to help the operator have situation awareness is to install many sensors with high capabilities (e.g., high-definition cameras on all of robot end-effectors). These sensors and their data, collected and presented to the operator, increase information regarding the remote environment and help the operator understand the remote situation. This further contributes to the operator's situation awareness with situational information (e.g., the environment's status, the robot's configuration, and the robot's relation to the environment). However, while an operator may develop some level of situation awareness using the data, it may also result in a negative outcome because of human limitation. People can pay attention to a certain amount of information simultaneously (Endsley, 2016) – that is, the robot may be sending too much information for the operator to understand in real-time.

Another method is to offload the operator's tasks to autonomous robots to focus on understanding the remote situation. High levels of automation are vital to assist teleoperators in managing their tasks and robots (Squire & Parasuraman, 2010; C. Y. Wong & Seet, 2017). However, even using advanced autonomous robots the operators encounter situations where they must control robots manually (e.g., deciding a course of actions, making a moral decision, or fixing inflexible predefined behaviors, Rosenfeld, Noa, Maksimov, & Kraus, 2016; Wright, Chen, & Barnes, 2018). When a robot needs the operator's involvement, they must comprehend the remote situation using the robot's streamed data before issuing any commands. That is, we still need a method to help the operator understand the remote environment using the robot's sensor data through teleoperation interfaces.

Teleoperation interfaces take an essential role: they must present the remote data in

a meaningful way for operators to understand the data easily and quickly. We tackle ego-centric teleoperation interface challenges through identifying interaction problems and investigating novel interface visualization strategies. This is because the teleoperation interface is one of the variables impacting an operator's overall task performance (J. Chen, Glover, Li, & Yang, 2016; Rea et al., 2017; Seo, Rea, et al., 2017) and because the way of presenting the information on the interface can significantly influence the operator's situation awareness (Endsley, 2016).

Research has proven that good interface designs assist users by reducing their cognitive load (Bark, Tran, Fujimura, & Ng-Thow-Hing, 2014; Rea et al., 2017; Singh et al., 2013) and improving task performance (Al-Megren & Ruddle, 2016; Ferland, Pomerleau, Le Dinh, & Michaud, 2009; O'Keeffe, Ward, & Villing, 2016). This assistance was also proven when the amount of data presented was held constant (Baudisch & Rosenholtz, 2003; Borst, Suijkerbuijk, Mulder, & van Paassen, 2006; Gustafson, Baudisch, Gutwin, & Irani, 2008; Potter et al., 2009). These findings are also applicable to teleoperation interface designs (Singh et al., 2013). Teleoperators can accomplish their missions successfully (Cohn, Older, Arnold, & Neill, 2013), complete their tasks more quickly (Labonte, Boissy, & Michaud, 2010; Leeper, Hsiao, Ciocarlie, Takayama, & Gossow, 2012; Mast et al., 2014; Saakes et al., 2013; Singh et al., 2013), reduce their number of errors (J. Y. C. Chen, Haas, & Barnes, 2007; Gombolay, Bair, Huang, & Shah, 2017; Singh et al., 2013), use human resources efficiently (O'Keeffe et al., 2016), and reduce their cognitive load (Rea et al., 2017; Seo, Rea, et al., 2017). Improving teleoperation interfaces to increase task effectiveness and lower operator effort is an ongoing research challenge in human-robot interaction (Kruckel, Nolden, Ferrein, & Scholl, 2015; Singh et al., 2013). In this work, we aim to help

teleoperators accomplish their missions by keeping a high level of situation awareness of the robot and the remote environment while reducing their overall workload.

1.2. Investigating Novel Interface Strategies to Improve Teleoperation Performance and Experience

This research targets improving the operator's situation awareness through designing, prototyping, and testing robot teleoperation interfaces. Our main approach is to collect the remote information and present them to the operator via teleoperation interface in a way that they can comprehend in real-time. The field already understands the importance of interface visualization and its impact on teleoperation effectiveness. This thesis presents the breadth-based exploration to discuss the wide-ranging teleoperation problems and their improvements with our novel interface designs.

We have identified four different remote teleoperation interaction challenges by classifying them in numbers of robots and information varieties. In that sense, we have two variants on the numbers of robots: a robot and a group of robots. Additionally, we have two different information varieties (one kind of information or many kinds) in our explorations. We have visualized our four exploration in this categories in Figure 1-2.

Our breadth-based explorations in teleoperation where operators must parse overwhelming amounts of information include: (a) teleoperating a robot with multiple cameras to increase the information volume, (b) teleoperating a robot in a team while maintaining awareness of group members' relative locations for team cooperation, (c) teleoperating a team of robots and monitoring their status for the effective use of human resources, and (d)

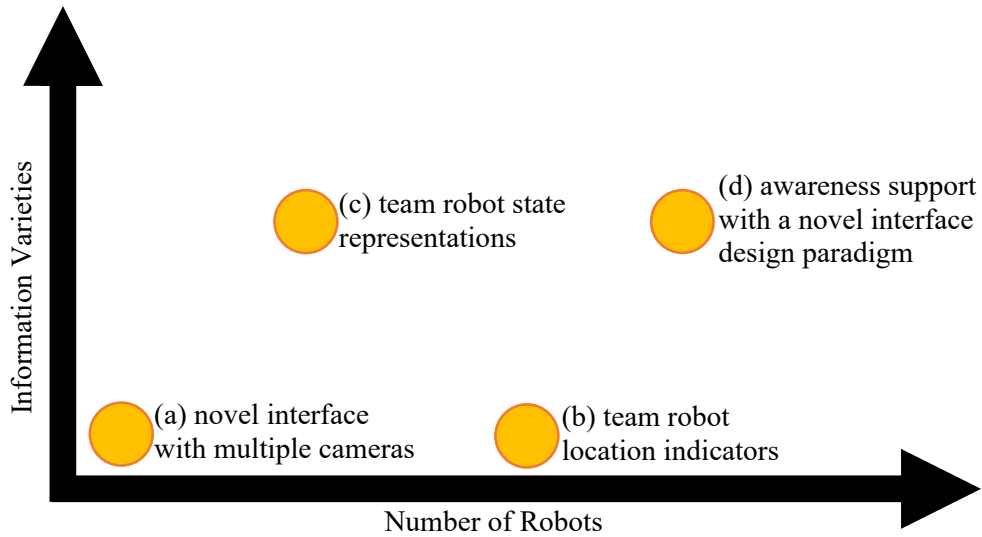


Figure 1-2. This research covers the breadth of the teleoperation challenges with the perspective of numbers of robots (horizontal) and information varieties (vertical). The alphabet letters are the order of our introduction in this document.

understanding pertinent teleoperation information in controlling multiple robots in turns.

The first problem we identified is teleoperating a robot with increasing numbers of camera feeds, especially with multiple specialized cameras. Multiple cameras are helpful to attain a high level of situation awareness; however, they simultaneously increase the teleoperator's workload. This increase is due to the operator's need to review all camera feeds to comprehend the information. We explore interface designs to help an operator understand multiple cameras' feeds while increasing their overall task performance.

Multi-robot teleoperation challenges go beyond the challenges of single-robot multi-camera teleoperation. The demand for an operator to teleoperate multiple robots increases to maximize mission efficiency (Lewis, 2013; C. Y. Wong & Seet, 2017), especially in exploration (e.g., deep-sea), military reconnaissance, or search and rescue (Zheng, Glas, Kanda, Ishiguro, & Hagita, 2013). In team robot teleoperation, team robot members' location information can help an operator maintain teamwork effectiveness. However, operators must maintain team location awareness in addition to all remote information. We

explore location indicators to support an operator's team robot location awareness with minimal impact to their overall task performance.

During team teleoperation, an operator must maintain awareness of team robot states (e.g., battery level) and the location information to ensure team effectiveness and coordination. However, it is unclear how to provide team robot information to an operator without providing a vast amount of information. We explore various information visualization representations and graphical visualization parameters to provide team robot member's states in egocentric teleoperation interfaces.

While controlling multiple robots, the operator must gain information of a specific robot. We propose a novel interface design paradigm that provides information during visual transitions between changes of robot control. This approach can help an operator focus on one robot under their control and have a high level of situation awareness of other robots, as necessary. In multi-robot teleoperation, an operator may face situations where other robots need the operator's direct attention to continue their work, such as making a moral decision. The operator must switch their control from one robot to another and understand the new robot's up-to-date situation. We call this process control transition. To accomplish tasks, the operator must understand various information (e.g., the robot's request, current situation, etc.). How quickly and how well they understand the information impacts their overall task performance. Thus, we provide information during the control transition.

In summary of the above introduction, we perform the following explorations:

- a) exploring a novel teleoperation interface to improve teleoperation effectiveness with multiple cameras;

- b) providing the location information of team robot members in egocentric teleoperation interfaces;
- c) exploring interface design techniques to represent multiple team robot states; and
- d) investigating how to support operator awareness in teleoperation control transition.

We target improving a teleoperator's situation awareness during the presentation of increased information, through novel teleoperation interfaces. In the following subsections, we introduce our investigations, approaches, and solution summary.

1.2.1. Improving Teleoperation Effectiveness with Multiple Cameras

Teleoperators control robots through unknown and dangerous terrain while monitoring multiple cameras' feeds from the remote environment (Kelly et al., 2010; Saakes et al., 2013). While multiple specialized views can improve task effectiveness (Hughes & Lewis, 2004), they increase an operator's mental demand to comprehend the larger volume of the information compared to a single feed. A cleverly designed teleoperation interface can minimize an operator's mental demand (e.g., Singh et al., 2013).

Currently, many teleoperation interfaces often use a toggle view or a tile view (Lewis, 2013), examples of which are presented in Figure 1-3 and Figure 1-4. Using these interfaces, an operator must remember the mode (which camera is currently on display) or continuously change their focus between views while understanding their physical relationship – depending on where they are pointing to, the camera perspective changes. In addition to the teleoperation difficulties (e.g., maintaining situation awareness), this mental overhead of the cameras negatively contributes to the operator's workload and impacts their overall task effectiveness (Drury, Scholtz, & Yanco, 2003; Mast et al., 2014; Singh



Figure 1-3. Canadian Army's unmanned ground vehicle (left) and controller box (right). The controller box contains camera buttons (yellow ellipse) on top of the robot diagram to switch camera feed. The monitor on the controller box displays one video feed at a time (i.e., toggle view).

Pictures are taken by Stela H. Seo, 2016.

et al., 2013).

Our approach presents multiple cameras' feed in a merged way that reduces the operator's mental overhead. That is, the interface merges multiple cameras' feeds into a single multi-resolution view using the detail-in-context design paradigm and displays them at once. With our detail-in-context approach (Figure 1-5), the operator does not need to change their focus nor perform a mental mapping between cameras.

We designed, built, and evaluated a novel egocentric teleoperation interface in a mock search and rescue scenario through a formal user study to understand our solution in an ecologically valid environment. From the formal study, we understood our solution's advantages and disadvantages against widely used solutions (base cases: toggling multiple views and putting views side-by-side). We found that our solution helped operators accomplish search and rescue tasks with fewer critical incidents than other interface solutions.

Our results highlight the importance of teleoperation interface designs in search and rescue with multiple cameras through our evaluation and analysis. Our novel teleoperation

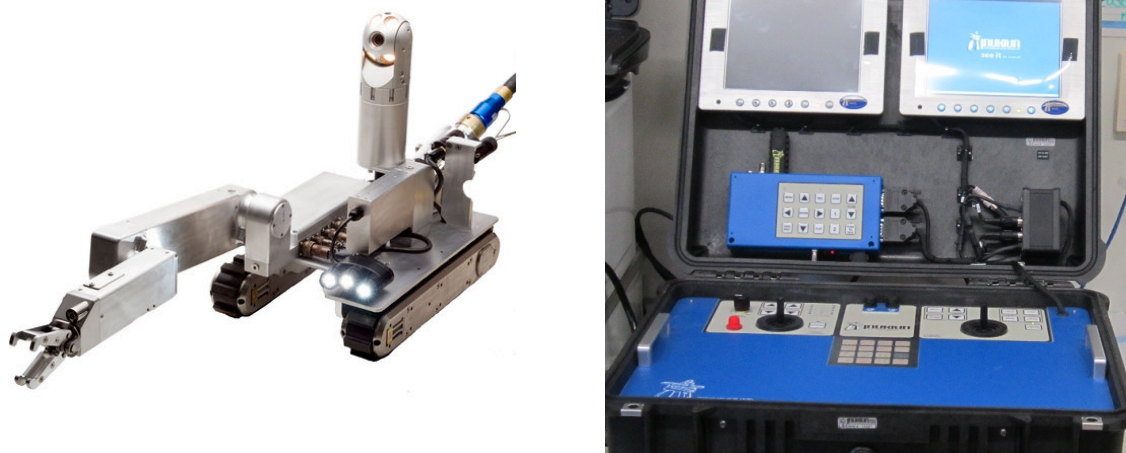


Figure 1-4. Versatrax 450 Robot Crawler (left) from Inuktun Inc. and its control box (right). The control box can display two camera feeds from the robot at once (side-by-side view).

Versatrax 450 photo was retrieved from inuktun.com on September 10, 2019. The control box picture was taken from the Human-Robot Interaction Lab, University of Manitoba, 2013.

interface helped an operator maintain a high level of situation awareness, reduce the operator's cognitive load, and improve their overall task performance.

1.2.2. Providing the Location Information of Team Robot Members

Teleoperation robots are increasingly working in teams (J. Y. Chen & Barnes, 2013; Glas, Kanda, Ishiguro, & Hagita, 2012; Lewis, 2013). In team teleoperation, awareness of team member locations is essential for successful and effective team coordination. As an example, it is critical to know where to go to assist a team robot (e.g., when the robot needs assistance to remove a large object) or where not to go (e.g., when a team robot is working on a task that may damage other nearby robots).

A standard solution for object locations in human-computer interaction and human-robot interaction is to include a top-down map (Glas et al., 2012; Recchiuto, Sgorbissa, & Zaccaria, 2016; Zheng et al., 2013), often called a mini-map, within the interface. Even when environmental data is not available, it is helpful to include the other robots' relative

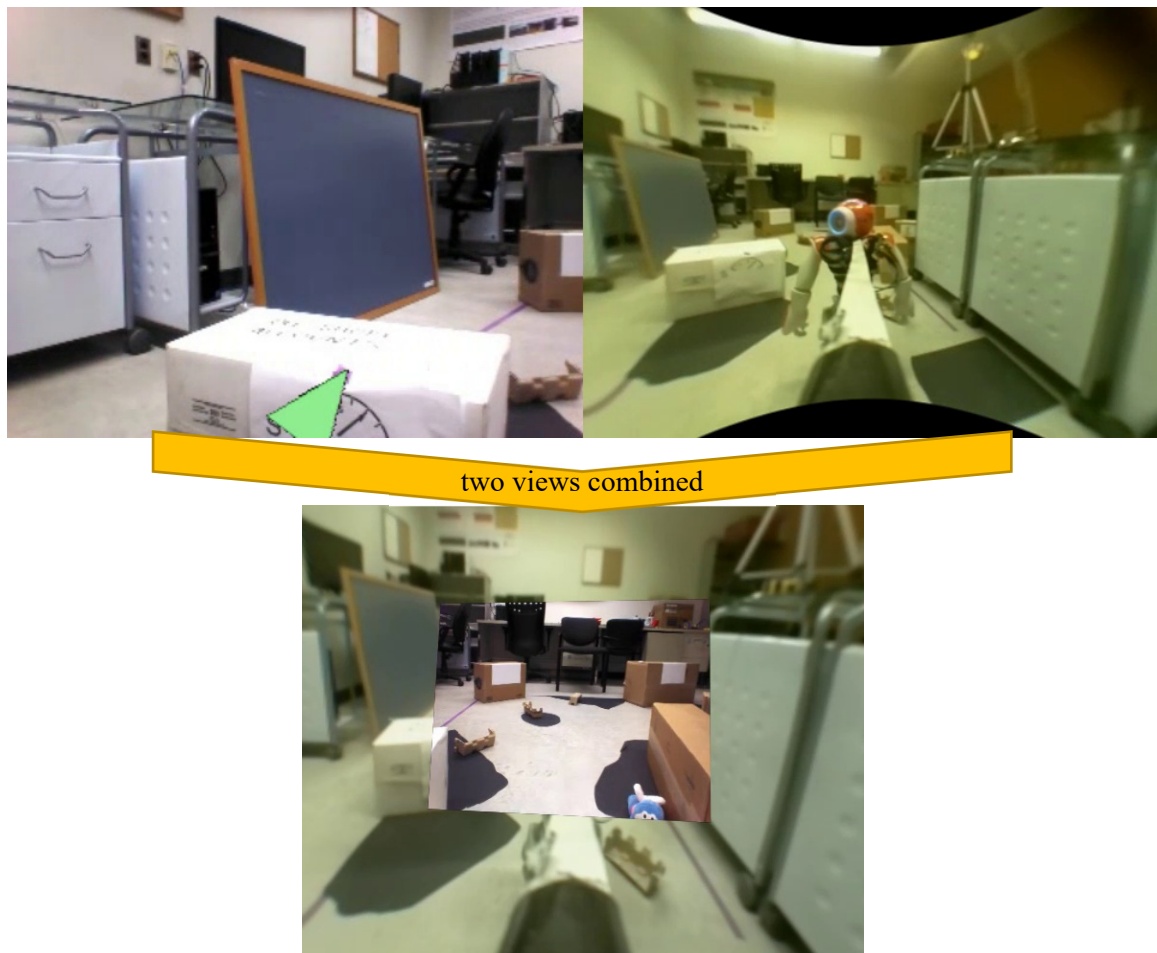


Figure 1-5. Our novel teleoperation interface (bottom) using two live camera feeds from a single robot. The interface merges two camera views of a single robot (top) into one mixed-resolution, detail-in-context view (bottom).

locations (Zheng et al., 2013). For example, Figure 1-6 has a mini-map on the right, which shows team robots' relative locations against the teleoperated robot (the black ellipse).

However, mini-maps have limitations. When they are presented with a camera feed, it either requires screen space when placed on the side (as in Figure 1-6) or hides a part of the video feed when overlaid. This is undesirable since the video has critical information for teleoperation (Herring et al., 2016; Jaju, Banerji, & Pal, 2013). Another implementation is to provide the map using a mode switch (i.e., toggle between the video feed or the map) or a separate monitor. With this implementation, an operator must switch their focus between the video feed and the map.

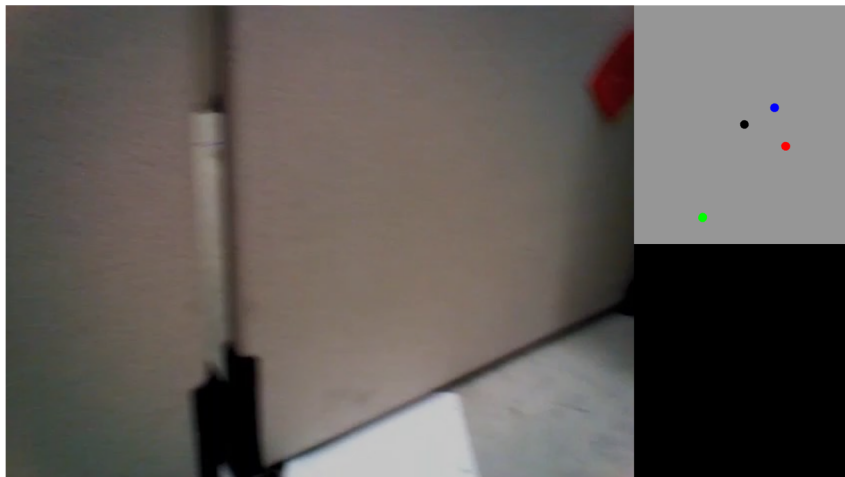


Figure 1-6. A mini-map example without the environmental data. It is on the remote video feed's right. There are four robots in the scene (black, red, blue, green ellipses).

We explore how to provide the team robot's location information to an operator without changing their focus and taking much screen space. In our scenario, the purpose of knowing the team robot's location is to move toward and support a team robot proactively, or to avoid any unnecessary contact for team cohesion (e.g., search wider area).

Our technique uses the operator's peripheral vision to maintain sufficient awareness of team locations in real-time, using significantly less screen real estate than mini-maps (Figure 1-7). We encode direction and distance of team members into location indicators displayed on top of the video feed (hence, in-feed). We designed our in-feed location indicators using adapted off-screen object visualization techniques from mobile human-computer interaction (Baudisch & Rosenholtz, 2003; Burigat, Chittaro, & Gabrielli, 2006; Gustafson et al., 2008).

We conducted a formal study to understand our in-feed location indicators' benefits and effectiveness compared to a simple mini-map. Further, we extracted our solutions' advantages and disadvantages from the study. Our results contribute to designing in-feed team robot location indicators in teleoperation interfaces.

In summary, our in-feed location indicators take a small portion of the screen space while providing enough information to maintain a general awareness of team robot locations, increasing the operator's overall situation awareness in team teleoperation.



Figure 1-7. One of our in-feed location indicators. It takes significantly less screen space compared to a mini-map. The implementation detail is in Chapter 4.

1.2.3. Exploring Interface Design Techniques for Graphical Representations of Team Robot States

In team teleoperation, maintaining awareness of robot team members and their robotic states is essential for efficient and successful team coordination. For example, an operator can decide which team member needs their help or how to communicate with the team member. Team robot awareness becomes critical when an operator manages and monitors multiple robots' workflow in real-time (J. Y. C. Chen & Barnes, 2012b; Kidwell, Calhoun, Ruff, & Parasuraman, 2012; Rosenfeld, Agmon, Maksimov, & Kraus, 2017).

Team teleoperation occurs in different ways. Two extreme examples are to have multiple people control a range of robots to provide better coverage (e.g., flying robot, high speed, different sensors), or a single operator controls all robots to reduce the human cost. Alternatively, situations allow some combination of both (J. Y. C. Chen & Barnes, 2012b; Zheng et al., 2013). However, in all cases working with a team of robots increases the

amount of information for operators to monitor for team cohesion, requiring more cognitive effort. This additional cognitive load is a problem as user error remains a primary cause of mistakes in teleoperation scenarios (J. Y. C. Chen et al., 2007; Gombolay et al., 2017).

There is a range of potential approaches to teleoperating robot teams, such as overview interfaces that enable a single operator to manage multiple robots (Guo, Young, & Sharlin, 2009), multiple operators each using their first-person view (Demir et al., 2017), or a single operator switching between active robots or views as needed (Calhoun, Warfield, Wright, Spriggs, & Ruff, 2012; Draper et al., 2008). We address the typical case where an operator is fully controlling one non-autonomous robot in an egocentric teleoperation interface and must maintain awareness of other robots in the team (and their actions).

This situation poses a range of interaction challenges such as ongoing communication with human team members, managing sliding-scale or mixed-initiative designs (Gombolay et al., 2017), supporting an operator when they context switch from controlling one robot to another (Draper et al., 2008), and representing the states and actions of robot team members in an easy-to-understand fashion. This project focuses on the last component, the specific awareness problem of an operator needing to maintain real-time awareness of other robots' actions and states while fully controlling one robot (see Figure 1-8). Our results transfer to a range of teleoperation situations, regardless of how autonomous the other robots are or who operates them.

Our strategy to avoid overwhelming an operator in team teleoperation is to remove unnecessary details and present the high-level encoded data, thereby enabling a quick understanding of the team robots' states. We select applicable mobile robot states across

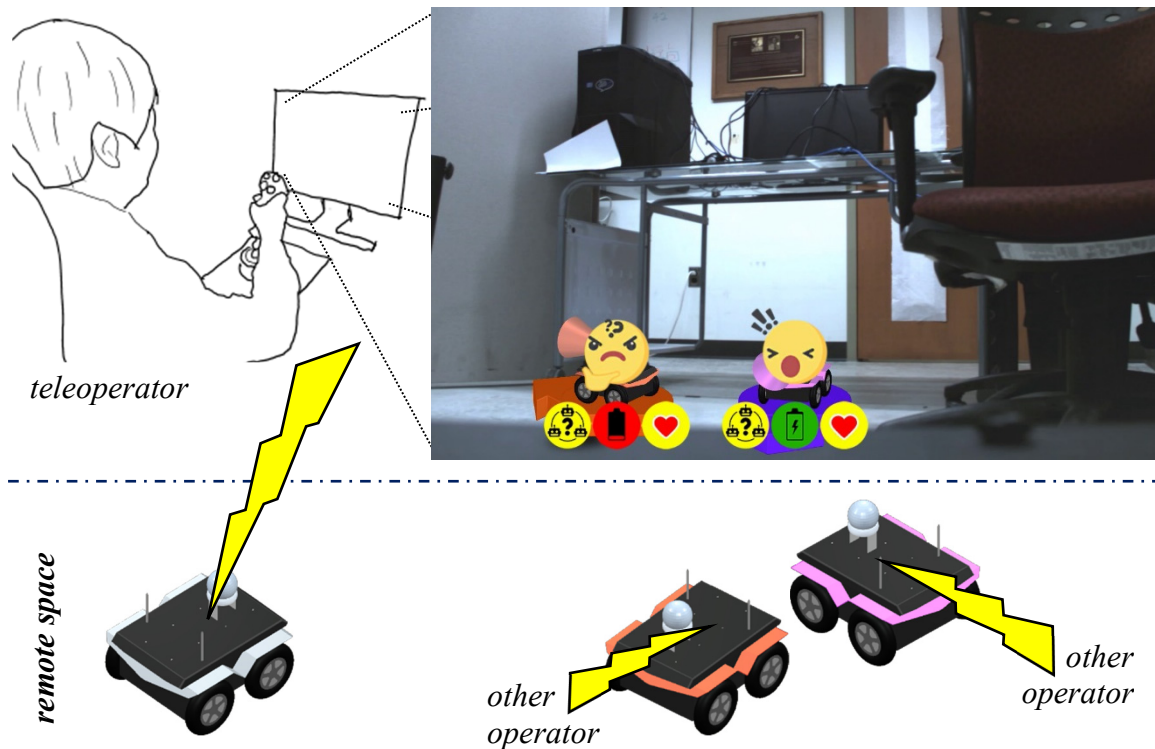


Figure 1-8. An operator teleoperates a robot located in remote environments, with two additional robot team members. While controlling one robot, they maintain awareness of other robots using on-screen widgets placed at the bottom of the screen.

search and rescue scenarios. Our goal is to generalize our graphical team robot state representations while having some specifics to be able to investigate use details in search and rescue scenarios. With the selection, we abstract details out to easy-to-understand status. We develop and explore a set of prototypes (Figure 1-9) with different interface parameters to simplify complex team member states, aiming to be easy for an operator to understand.

We designed a series of novel visualizations to represent states of other robots in a team, drawing from related work and human perception knowledge (Figure 1-8). Through a series of pilot studies, we conducted an iterative exploration with the general public. We collected people's opinions on different information representations and graphical visualization parameters for team robot state visualization. We present emerging themes for interface designers to design and present team robot state representations, helping an operator



Figure 1-9. Examples from our exploration space. We have three information representation techniques (text, icon, emotional information encoding) and graphical visualization parameters (color coding, animation, and the number of team robots). The detail can be found in Chapter 5.

maintain a high level of situation awareness in team teleoperation.

Our work yielded a set of novel widgets for graphically representing robot states, a testbed design for exploring teleoperation interfaces, and a set of initial guidelines for designing these interfaces. We further include a reflection on our evaluation methods with recommendations for ongoing work (see Chapter 5 for detail).

1.2.4. Informative Visual Transition in Multi-robot Teleoperation

One way to assist operators to control more robots is to increase robot autonomy. However, even the most advanced autonomous robots still need the human operator's involvement, such as when it encounters unexpected circumstances (e.g., immobilized in the middle of operation, Rosenfeld et al., 2016), or needs to make critical final decisions (Wright et al.,

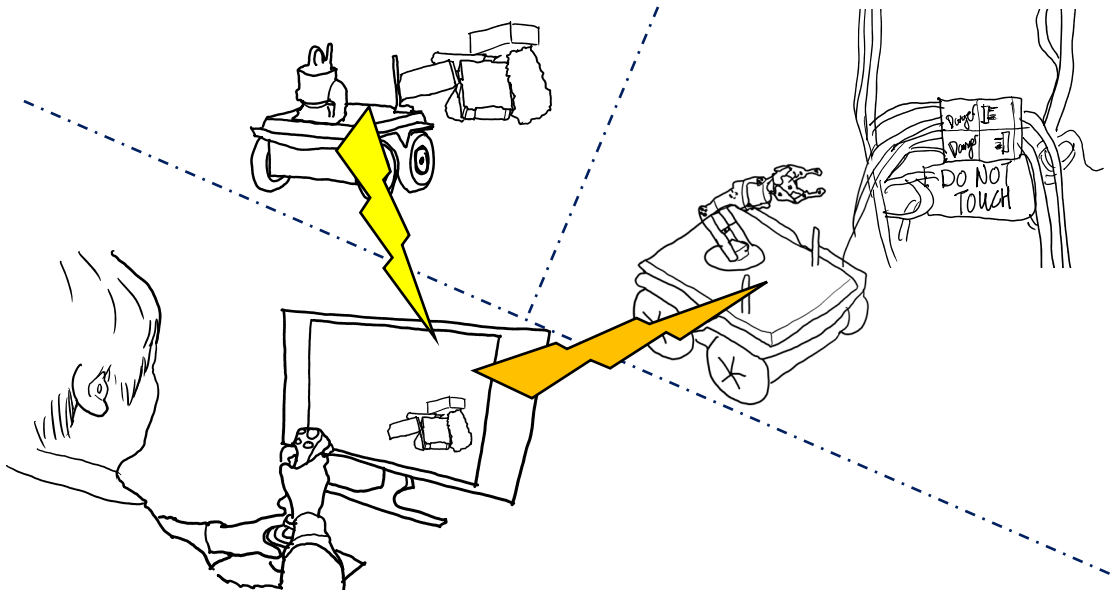


Figure 1-10. An operator remotely controls a robot to avoid debris (left); meanwhile, another robot (right) requests an operator's attention to work on a complex object. To help the robot on the right, first the operator must understand the situation around the robot.

2018). In this situation, for an operator to take the control of the robot and send an appropriate command to resolve the situation, the operator must switch their focus and control from their current task to the new robot in need (e.g., controlling a different robot). We call this *control transition*.

Control transition from one robot (or task) to another is a cognitively expensive task. During a control transition, the operator must discard the information regarding their current task, and acquire and understand the information regarding the robot in need. For example, in Figure 1-10, an operator, who is navigating one robot, receives a help call for another robot to require their attention; they must switch their focus to that robot and survey the situation before giving any commands. That is, the operator must achieve a high level of situation awareness (i.e., understanding the in-need robot's current states, what it was doing, where it is in its environment, and what else is in the environment Endsley, 2015) to appropriately operate the new robot.



Figure 1-11. A video game's loading screen with additional background information. *Dragon Age: Inquisition*, Electronic Arts, 2014.

Our goal is to support the operator's situation awareness during a control transition to a new robot by presenting relevant information. We call this *informative visual transition*, as we provide information during the interface's visual transition (as a part

of the control transition). We recognized a similar problem in film and other medium. They must orient audiences to a new scene and situation to ensure that the audiences can follow the story arc. For example, movies use a slow pan or dissolving effects of two scenes to provide a transitional effect (Katz, 2019), novels use narrative exposition as a scene transition (Hill, 2016), and video games use a loading screen to provide additional information (Figure 1-11).

We drew on techniques from literature in film and media and discussed with human-robot interaction experts to develop a teleoperation design framework for informative visual transitions. The framework's goal is to provide example techniques from other relevant fields (e.g., film, video games) and have a vocabulary and keywords that explain how and why they are helpful in teleoperation. With the framework, we enable the further discussion in designing informative visual transitions in teleoperation.

1.2.5. Summary

In teleoperation, the operator needs to achieve a high level of situation awareness of the remote environment. It is extremely important especially in search and rescue where the

operator must make crucial decisions within a limited time frame. To support the operator, the interface (the only gateway) must present the remote information in understandable ways in real-time. We take the breadth-based approach to explore a variety of novel teleoperation interface designs and discuss our explorations in the following chapters.

1.3. Contributions

For successful teleoperation, an operator needs to maintain high situation awareness of robots, remote environments, and their missions. A cleverly designed interface can help an operator understand the teleoperation information easily and quickly, increase the operator's situation awareness, and increase their overall task performance, while not impacting their cognitive load (Endsley, 2016; Rea, Seo, & Young, 2020; Yanco & Drury, 2004b).

This thesis begins by addressing the identified challenges in search and rescue teleoperation interfaces. Following this, we investigated novel interface techniques and presented and discussed our teleoperation interface designs. Subsequently, we tested our novel teleoperation interfaces, and we can show that the interfaces help the operators develop a high level of situation awareness while reducing negative impacts on their cognitive load in teleoperating physically separated robots. Finally, our novel teleoperation interfaces can enable the design discussions in building egocentric teleoperation interfaces. Our findings can help many people in teleoperation fields, including search and rescue crews working in disaster zones, caregivers working with seniors remotely, and laypeople who cannot physically go to places.

In summary, our contributions are

- a novel interface design, prototype, and evaluation for simplifying interfaces that

have multiple camera views into an environment;

- novel in-feed location indicators for team robot members to maintain team location awareness in a hectic team teleoperation scenario;
- design explorations to present team robot members' working states through a series of user studies; and
- a novel design framework inspired by film and media techniques to achieve informative visual transition in egocentric multi-robot teleoperation interfaces.

Overall, this work will improve egocentric teleoperation interface designs by attacking pertinent problems facing teleoperation. We have created a range of novel interfaces throughout this work and developed knowledge to support teleoperators in different scenarios. Each project will eventually contribute to the more extensive and ongoing discussions on teleoperation interface design elements and shape future directions for in-depth exploration.

Chapter 2 Background and Related Work

Teleoperation is an operation of a machine or a robot at a distance using electric signals. The history of teleoperation goes back to the late 19th-century (613809, 1898). Teleoperation has always had a considerable influence on the military (e.g., military drones¹), entertainment (e.g., remote-controlled cars²), and other human-support applications (e.g., telepresence and search and rescue robots).

The concept of teleoperation is simple: retrieving information from a remote machine and sending control commands to the machine. However, teleoperation is challenging and cognitively taxing (Rea et al., 2017; Seo, Rea, et al., 2017). The operator must monitor a large volume of remote information, understand them, and make real-time decisions. That is, the operator must achieve and maintain their *situation awareness*. With a high level of situation awareness, an operator can be aware of their robot, mission, and the environment and plan their actions accordingly.

The operator's situation awareness is affected by system capability (of the control station or the remote robot), device connectivity, network delay, and teleoperation interface (Endsley, 1995). This variety of effects creates different teleoperation challenges. We focus our discussion on teleoperation interfaces and their designs.

To further discuss situation awareness, let's take a simple teleoperation example: to win a race using a remote-controlled car, the operator must have the information of their car, other's cars, and the track (e.g., their car's control latency, speed, and location; rivals' locations; potholes in the environment; etc.), understand the information (e.g., how far their

¹ en.wikipedia.org/wiki/History_of_unmanned_combat_aerial_vehicles

² rcroundup.com/radio-controlled/rc-cars/history-of-rc-cars/

car is from rivals), project the future states (e.g., keeping this speed at the corner may result in a severe collision), and make their decisions based on the information and their understanding. That is, to win a race, the operator must have a high level of situation awareness and make informed decisions. With sophisticated machines such as robots, a teleoperator has far more data to comprehend for their tasks and a complex control system.

Teleoperation interfaces have an essential role in presenting the necessary information (e.g., sensor data, robot states, and environment information) in a meaningful way and transferring the operator's commands swiftly to the remote device. For teleoperating remote robots, the interface must be able to steadily present the remote information retrieved from multiple sensors on the robots, support the operator to achieve a high level of situation awareness, and complete cognitively taxing control tasks (Nielsen et al., 2007; Parasuraman, Caccamo, Fredrik, Ogren, & Neerincx, 2017; Yanco & Drury, 2004b).

Improving interface design and usability is an active research focus in human-computer interaction and human-robot interaction. There is a body of work on improving teleoperation interface design and usability (Rea et al., 2017). Improving teleoperation interfaces is closely related to improving the operator's overall task performance (Endsley, 2011; Yanco & Drury, 2004b). This work presents novel teleoperation interface designs and discusses improvements in teleoperation interfaces. We tackle a broad spectrum of different teleoperation challenges. Before discussing our novel interface designs in-depth, we discuss a fundamental problem in teleoperation and other improvements in this section.

2.1. Situation Awareness, Teleoperation Challenge

To achieve teleoperation goals successfully and efficiently, an operator must have a high

level of situation awareness.

“Situation awareness is the perception of elements in the environment within a volume of time and space (level 1), the comprehension of their meaning (level 2), and the projection of their status in the near future (level 3)” (Endsley, 1995).

Situation awareness is an active concept in many domains, including aviation psychology (Durso & Gronlund, 1999), automobile driving (Mao et al., 2018), and robot teleoperation (Rea & Young, 2018; Son et al., 2013; Yanco & Drury, 2004b). The term, situation awareness, emerged from aviation psychology to describe the pilot’s understanding of tactical flight operations (Durso & Gronlund, 1999). However, situation awareness is not limited to aviation psychology: it is relevant to any cognitive activity and information processing, including teleoperation (Durso & Gronlund, 1999; Endsley, 2015; Yanco & Drury, 2004b). The critical idea of situation awareness in those domains is to explain the operator’s understandability of the current situation and project the future. The operator’s situation awareness leads them to make informed decisions, which impact their task overall performance (Endsley, 2015; Yanco & Drury, 2004b).

Researchers studying situation awareness focus on providing a foundation for improving system designs (Endsley, 1995). The foundation explains various components that influence an operator’s situation awareness, including interface designs – the influential aspects of situation awareness are summarized in Figure 2-1 (Endsley, 1995).

Situation awareness can explain a person’s cognitive processing (Durso & Gronlund, 1999). For example, for a person to accomplish even a simple goal like making a room bright, they need to understand current states, like their lamp (e.g., is it on or off,

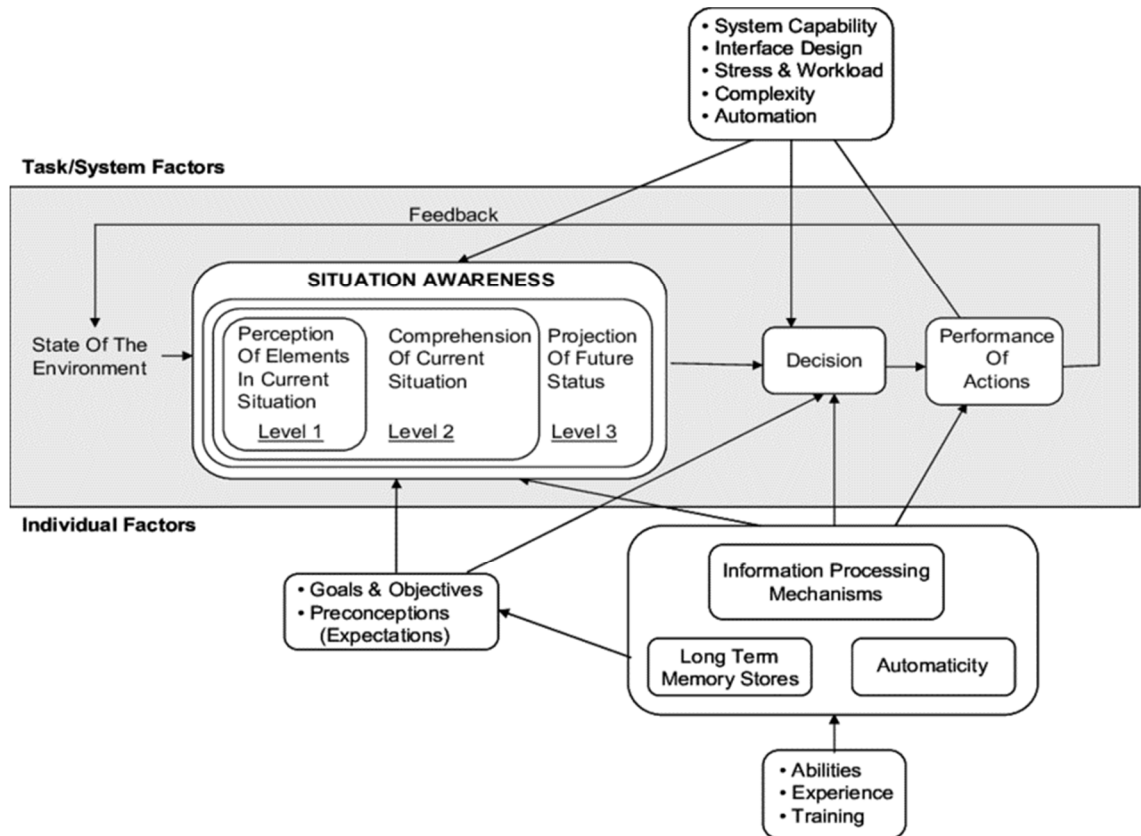


Figure 2-1. Model of Situation Awareness in dynamic decision making (Endsley, 1995). The figure is from the publication.

can you change the level of luminance, or can you change where it faces), their options (e.g., do you have other lamps, do you have a window, or is it bright outside), possible future states (e.g., the room may get bright anyway by the sunshine through the window), and other variables (e.g., is the lamp covered, or is it working or broken). After understanding the information (i.e., achieving situation awareness), the person can decide what to do to accomplish their goal. The consequences of their action(s) alter the environment (e.g., turn on the switch to illuminate the room), and the person iterates this process until their goal is achieved. Even in this simple example, there is an abundance of cognitive processing done by the person. Compared to the above simple example, teleoperation is far more challenging as there is much more information for the operator.

For robot teleoperation, a high level of situation awareness of an operator is the understanding of various information including their mission goals, internal and external states of their robots (e.g., physical configuration or battery level), the dynamic environment, and the projection of the robot's future behaviors (Yanco & Drury, 2004a). Different approaches support the teleoperator's situation awareness and depend on whether the teleoperation is collocated or non-collocated (Goodrich & Schultz, 2007; Lewis, 2013). During collocated teleoperation (i.e., proximate interaction), the operator can get environment information without an interface and display device to show additional information (e.g., head-mounted augmented reality display or a handheld device, Barber et al., 2015).

On the contrary, in remote teleoperation, it is not simple to relay a perception of elements from the environment (i.e., the level-1 situation awareness, Figure 2-1). As the operator is not in the same environment as the robot, the operator's option to build their perception is limited to relying on the teleoperation interface. While the operator struggles to build their perception, the remote robot tries to help the operator by sending as much information as possible (Liu & Nejat, 2013). However, the sensor capabilities, the network bandwidth, and the unexpected disconnection limit the information flow (Julio & Bastos, 2015; Parasuraman et al., 2017). It also takes time for the operator to iterate through the information they received from the remote robot. People have limited individual capacities to pay attention to simultaneous information (Endsley, 2016). Hence, the overall situation awareness reduced and the difficulty of having a high level of situation awareness increases in remote teleoperation.

Interface designers suggested improving remote information visualization to help the operator perceive more information at once (left in Figure 2-2, Seo, Rea, et al., 2017).



Figure 2-2. Two examples of improving visualization on teleoperation interfaces. The left interface integrates two cameras' feeds to show more information to the operator (Seo, Rea, et al., 2017). The right interface augments extra visual objects to help the operator locate the critical information quickly (Rea et al., 2017).

In addition to providing more information at once, improved interface designs can help the operator increase their comprehension of available information (right in Figure 2-2, J. Y. C. Chen et al., 2007; Gombolay et al., 2017; Rea et al., 2017), or even have the future states of the robot by projecting expected outcomes (Singh et al., 2013).

Improving interface designs helps the operator achieve a high level of situation awareness and increases their performance (Endsley, 1995; Rea et al., 2017; Seo, Rea, et al., 2017). Exploring interface designs to increase the operator's situation awareness is still one of the primary challenges. This thesis attempted to advance teleoperation interface designs by introducing novel interface designs and their potentials. Eventually, we aim to explore interface designs that help teleoperators have a high level of situation awareness.

2.2. From Human-computer Interaction and Human-robot

Interaction to Novel Teleoperation Interfaces

Human-computer interaction focuses on the interaction between people and computational

machines (Hewett et al., 1992).

“Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them” (Hewett et al., 1992).

To design and engineer an interactive computing system that creates practical user experiences, “designers need to know many different things about users, technologies, and interactions between them” (Rogers, Sharp, & Preece, 2007). This focus leads to human-computer interaction having many connections with other disciplines. For example, designing a social robot interaction to make people empathize with the robot requires a robotics and psychology background (Seo, Geiskkovitch, Nakane, King, & Young, 2015).

An interactive computing system consists of a user interface. Every time a person uses a machine, it is an interface that informs and records their activities, such as entering a command or receiving feedback (Rogers et al., 2007). There is a body of work to visualize different data in user interfaces such as map data (Brosz, Carpendale, & Nacenta, 2011), weather predictions (Potter et al., 2009), texts (Bongshin Lee, Riche, Karlson, & Carpendale, 2010), optical images (Gregori, Lam, Gregori, & Ranganathan, 2012), graph data (Collins, Penn, & Carpendale, 2009; N. Wong, Carpendale, & Greenberg, 2005), motions (Pingali, Opalach, Jean, & Carlbom, 2001), videos (Li, 2014; Meghdadi & Irani, 2013). Researchers continue to explore interface techniques to help increase users’ task effectiveness (Dziubak, Dubois, Bunt, & Terry, 2016; Ens, Finnegan, & Irani, 2014; Htun, Halvey, & Baillie, 2017; Seo, Rea, et al., 2017), improve situation awareness (Riek, 2012; Yanco & Drury, 2004b), and reduce task completion time (Komiyama, Miyaki, &

Rekimoto, 2017; Seo, Rea, et al., 2017; Singh et al., 2013). A common theme is helping users understand the data easily and quickly while increasing their task performance.

The principles of human-computer interaction and their work to explore interface techniques extend to human-robot interaction (J. Y. C. Chen et al., 2007; O’Keeffe et al., 2016). In human-robot interaction, many researchers explore different kinds of interactions between people and robots such as robot embodiments (Seo et al., 2015; Zheng, Glas, Kanda, & Ishiguro, 2012), motions to present a robot’s intention (Seo, Takashima, Young, & Kitamura, 2014; Sharma, Hildebrandt, Newman, Young, & Eskicioglu, 2013; Singh & Young, 2013), or teleoperating a robot remotely (Khan, Fitzmaurice, Almeida, Burtnyk, & Kurtenbach, 2004; Rea et al., 2017; Seo, Rea, et al., 2017).

Researchers continue to present design guidelines and principles for interfaces in various applications in both human-computer interaction and human-robot interaction. The motivations for this research include improving usability (Gittins, 1986; Gustafson et al., 2008; Kobsa, Koenemann, & Pohl, 2001), enhancing user experience (Hashimoto, Ishida, Inami, & Igarash, 2011; Labonte et al., 2010; Richer & Drury, 2006), reducing people’s workload (Chien et al., 2018; Lewis, 2013; Seo, Rea, et al., 2017; Singh et al., 2013), and increasing task performance (Labonte et al., 2010; Leeper et al., 2012; Mast et al., 2014; Saakes et al., 2013; Singh et al., 2013).

These goals also apply to improving teleoperation interfaces and help the operator accomplish their mission (Yanco & Drury, 2004a). Following this trend, our teleoperation interfaces help an operator maintain high situation awareness and improve overall task performance. In our work on designing teleoperation interfaces, we use different interface

designs for different challenges. The next section introduces techniques from human-computer interaction and human-robot interaction that we have considered or used for our novel teleoperation interface designs.

2.2.1. Improving Teleoperation Interface Designs

Robot teleoperation is a form of human-robot teaming, also known as human-agent teaming. The word agent refers to an intelligent software agent, an autonomous robot, or a teleoperated robot. Frequently, assumptions are made regarding common knowledge, shared understanding of objectives, and two-way communication between a person and a group of robots to address challenges in human-robot team scenarios (J. Y. Chen & Barnes, 2013; Lewis, 2013). However, during real teleoperation scenarios, the assumptions might be violated at any time due to unexpected situations.

In many teleoperation fields, operators control and manage more robots to efficiently use human resources (J. Y. Chen & Barnes, 2013; Glas et al., 2012; Lewis, 2013). Multi-robot teleoperation interfaces play a critical role in accomplishing mission by connecting a person (operator) with a remote group. The interface is a crucial factor for an operator's task performance (J. Chen et al., 2016; Rea et al., 2017; Seo, Rea, et al., 2017) and situation awareness (Phillips & Jentsch, 2017; Seo, Rea, et al., 2017). This also applies in multi-robot teleoperation. Thus, teleoperation interfaces are important to help the operator maintain their awareness of mission goals, remote environments, and mobile robots.

Improving interface usability is an ongoing challenge for UAV control (Richer & Drury, 2006), inspection (Singh et al., 2013), domestic robots (Labonte et al., 2010; Mast

et al., 2014), or medical consulting (Agarwal et al., 2007). Techniques to improve the usability of teleoperation include novel control schemes (e.g., Hashimoto et al., 2011; Saakes et al., 2013; Sakamoto et al., 2009) that abstract away low-level control problems to provide high-level controls of the robot (Leeper et al., 2012; Sakamoto et al., 2009; Singh et al., 2013), or improved data presentation (Drury et al., 2003; Hashimoto et al., 2011; Labonte et al., 2010; Richer & Drury, 2006). Many come down to improve task performance such as completion time (Labonte et al., 2010; Leeper et al., 2012; Mast et al., 2014; Saakes et al., 2013; Singh et al., 2013), to reduce teleoperator errors (J. Y. C. Chen et al., 2007; Gombolay et al., 2017), to use human resources efficiently (O’Keeffe et al., 2016), and to reduce an operator’s cognitive load (Rea et al., 2017; Seo, Rea, et al., 2017). Our work focuses on improving teleoperation interface designs to help operators achieve a high level of situation awareness, improve their overall task performance, and human errors. We use successful methods such as improving data presentation or abstract details out.

A common design pattern for teleoperation interfaces in search and rescue is video-centric visualization which builds an interface around the remote video feeds (e.g., Figure 2-3, Herring et al., 2016; Jaju et al., 2013; Singh et al., 2013). The video feeds contain crucial information for the operator to understanding the remote environments (Herring et al., 2016; Jaju et al., 2013; Singh et al., 2013). The bigger screen resolution helps an operator understand the feeds easily (Seo, Rea, et al., 2017). Therefore, it is essential to design teleoperation interfaces to keep the video at the maximum size and reduce overlaps (Herring et al., 2016; Jaju et al., 2013; Singh et al., 2013). In this work, we design our novel teleoperation interfaces with the video feed (main information source) at the maximum size while providing additional necessary information to the operator.

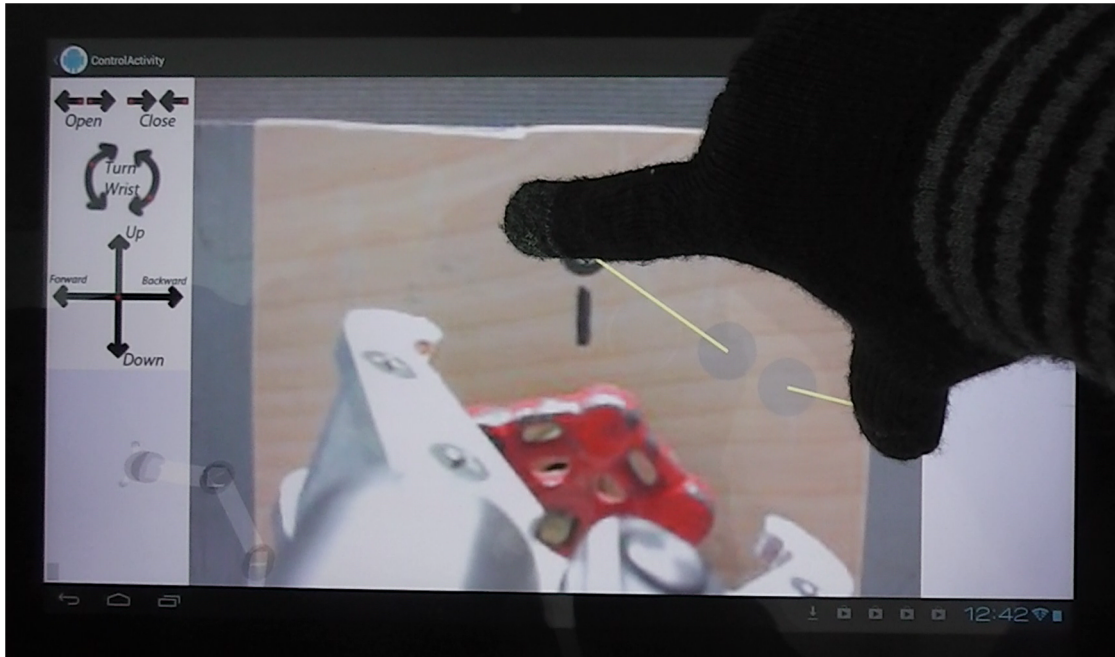


Figure 2-3. An example of video centric teleoperation interfaces (Singh et al., 2013). The interface primarily has a remote video feed on the screen and has other information around the video feed.

Graphics are commonly embedded into video, as flat or augmented overlays to maximize video size and support an operator focusing on the feed (Jia et al., 2015; D. Lee et al., 2013). Another method is to use the display's periphery to minimize overlaps (Baudisch & Rosenholtz, 2003; Burigat et al., 2006; Gustafson et al., 2008). Such ecological interface design has been used to support the mapping and situation awareness required for particular tasks, for example, fusing sensor data into single task-oriented displays and widgets (Borst et al., 2006). Camera-specific work has embedded a single video feed within virtual environments to show a robot on a 2D map or within a 3D environment, and projecting the camera's view at the correct angle and location within the scene (Labonte et al., 2010; Ochiai, Takemura, Ikeda, Takamatsu, & Ogasawara, 2014; Saakes et al., 2013).

Interface designers and researchers use a variety of methods to reduce an operator's cognitive load. These methods include reducing the data complexity (Seo, Young, & Irani, 2017), visualizing high-level task-relevant information (Yang, Kamezaki, Sato, Iwata, &

Sugano, 2015), and highlighting key locations to reduce the chance of missing an important event (Rea et al., 2017). Our work has a similar theme: we explore novel egocentric teleoperation interface designs to help an operator maintain a high level of situation awareness and increase their overall task performance.

2.2.2. Multi-video Stream in Egocentric Video Teleoperation Interfaces

In search and rescue scenarios, one predominant control paradigm is for an operator to stay in a safe area, watch video feeds, and send commands to the robot remotely using egocentric teleoperation interfaces (Casper & Murphy, 2003; Liu & Nejat, 2013; Nielsen et al., 2007; Singh et al., 2013). Egocentric interfaces allow an operator to control the robot as if they are in or with the robot. For example, when an operator controls a robot to move forward, every onboard sensor moves together with the robot and the information (e.g., video feed, GPS) changes relatively. This type of interface can provide an operator with the feeling of *being in the robot*; the video feed is often a first-person view, but it can have the robot in it (Saakes et al., 2013).

One of the challenges with egocentric video teleoperation interfaces is to provide an operator with multiple video streams. Teleoperators control robots with multiple cameras (Figure 1-3, p.11) while monitoring multiple camera feeds into the remote environment (Kelly et al., 2010; Saakes et al., 2013). The multiple cameras (or increasing amount of information) can help an operator's situation awareness up to a certain level; however, they also increase operator cognitive load (Drury et al., 2003; Kelly et al., 2010; Yanco & Drury, 2004b).

Successful teleoperation requires the operator to have a high level of situation

awareness (as discussed earlier in this chapter), which requires them to understand various teleoperation-related information. We explore novel egocentric video teleoperation interfaces to help an operator obtain the given information quickly while reducing their cognitive load. Our exploration's detail is in Chapter 3, Multiple Cameras: Interactive Detail-in-context Interface Using Multiple Cameras in Teleoperation.

2.2.3. Group Awareness in Multi-robot Teleoperation

Teleoperation robots, especially for search and rescue missions, are increasingly working in teams (J. Y. Chen & Barnes, 2013; Glas et al., 2012; Lewis, 2013) to increase the search area, to quickly respond to distress calls, and to extract as much remote information as possible. In previous work in human-agent interaction, researchers found that good teamwork involves proactive actions to help team members (Demir et al., 2017) and positive engagements between team members (Price & LaFiandra, 2017). The system must help a team-member operator be aware of others' states, helping them maintain high situation awareness and support proactive actions.

Multi-robot teleoperation increases its difficulty when a team of robots deployed in a scene are not in the same task environment. When team members are out of sight, beyond the remote teleoperation difficulties, being aware of team-member working states introduces additional challenges. The operator must rely on the information retrieved from multiple robots for team cohesion and their robots. Thus, the challenge to maintain a high level of situation awareness exacerbates in team teleoperation (J. Y. Chen & Barnes, 2013; Glas et al., 2012; Lewis, 2013).

There is a body of work in multi-robot teleoperation; for example, researchers study

how to work with a large number of robots (J. Y. Chen & Barnes, 2013; Kolling, Walker, Chakraborty, Sycara, & Lewis, 2016), provide commands to teams effectively (Glas et al., 2012; D. Lee et al., 2013; Lewis, 2013), design visual interfaces (Kolling et al., 2016; Lewis, 2013; Recchiuto et al., 2016; Zheng et al., 2013), or provide increased amounts of state or sensor data effectively (Glas et al., 2012; Hong, Lee, Bühlhoff, & Son, 2016; Lewis, 2013). Further studies include interfaces for controlling multiple robots from a single, meta interface (Hong et al., 2016) and controlling a single robot at a time as a part of a large team (J. Y. Chen & Barnes, 2013; Glas et al., 2012). A common goal among these research endeavors is to increase the information available while minimizing the operator's cognitive load. Our work follows this endeavor to maximize information available for the operator's situation awareness, while reducing their cognitive load by abstracting unnecessary details out and by improving data visualization.

For controlling or monitoring multiple robots, some use tile or picture-in-picture views (Lewis, 2013; Richer & Drury, 2006), with other information (e.g., a map, Ion, Chang, Haller, Hancock, & Scott, 2013) to support location awareness (Zheng et al., 2013). Many researchers worked on improving artificial intelligence for mobile robots (Rosenfeld et al., 2017), control mechanisms of teleoperated robots (Hayes & Shah, 2017; D. Lee et al., 2013; Omidshafiei et al., 2017), and visual interfaces (Rea et al., 2017; Seo, Rea, et al., 2017). Team robots in our explorations can either be automated or controlled by other human operators. We focus on teleoperation interface designs, instead of team robots' control mechanisms, to support the operator's understandability of the situation.

Due to the increasing amount of information retrieved from remote robots, multi-

robot teleoperation interfaces must present more information than single-robot teleoperation interfaces. One way to reduce the volume of the data is to abstract details away and design simple visualization. In human-computer interaction and human-robot interaction, researchers explored iconic representations (Bartram, Ware, & Calvert, 2003; Gittins, 1986; Selkowitz, Lakhmani, & Chen, 2017), emotional encodings (Paul & Komlodi, 2012; Seo et al., 2015; Sharma et al., 2013), and color encodings (Murch, 1985) in several applications. Icons can metaphorically convey a complicated message (Gittins, 1986) and can be understandable at a distance (Kline, Ghali, Kline, & Brown, 1990) and while moving (Long & Kearns, 1996). Emotional encoding with a single facial expression “may well serve as both symptoms of an underlying state and communicative signals” (Hess & Hareli, 2015; Parkinson, 2005). Color coding can enhance an interface’s effectiveness (Murch, 1985). The number of metaphoric visualization techniques in these fields proves their usefulness. We also explore their feasibility of improving situation awareness and task performance in multi-robot teleoperation.

We focus on addressing the challenge of maintaining awareness of team member states using different interface design techniques. These techniques include leveraging peripheral vision with abstract information (Chapter 4), encoding team states into visual widgets (Chapter 5), and providing the information only when it is necessary (i.e., switching control to a robot, Chapter 6).

2.3. Summary

This chapter discusses the definition of teleoperation, the challenge of situation awareness, and previous work to improve teleoperation interface designs. Our primary goal is to help

an operator increase their situation awareness to respond to any changes and urgent situations during their mission.

We discussed teleoperation interface research and visual interface techniques in the fields. The body of work in the field shows that good interface designs can help users including teleoperators maintain situation awareness and improve their overall task performance. However, the work remains challenges in real-time teleoperation scenarios.

Our focus is to help an operator control a robot with multiple visual data sources (i.e., cameras) or increase team cohesion in a teleoperation team with high situation awareness. Primarily, we focus on helping the operator quickly and easily understand the visual teleoperation information. We explore problems of understandability and their solutions through multiple projects. We merge all lessons from them and create design strategies for multi-source teleoperation interfaces. While there are many applications in teleoperation with multiple data sources, we carefully choose selective cases to cover broad varieties.

In the following chapters, we describe the challenges of search and rescue scenarios in detail, our approach to address the challenges, and our findings from this work. Finally, we discuss how our general approach of improving interface design can impact an operator's situation awareness and overall task performance and suggest future work from our findings, limitations, and lessons learnt throughout the process.

Chapter 3 Multiple Cameras: Interactive Detail- in-context Interface Using Multiple Cameras in Teleoperation

We published a peer reviewed paper on this work and gave a public presentation at a conference. Some ideas and figures in this chapter are from that paper.

Stela H. Seo, Daniel J. Rea, Joel Wiebe, and James E. Young. 2017. Monocle: Interactive detail-in-context using two pan-and-tilt cameras to improve teleoperation effectiveness. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 962–967. doi.org/10.1109/ROMAN.2017.8172419

Search and rescue teleoperators are typically near the disaster site, watch video feeds, and provide commands to a mobile robot using an egocentric video teleoperation interface (Casper & Murphy, 2003; Liu & Nejat, 2013; Nielsen et al., 2007; Singh et al., 2013). Using the video feed, the operator finds victims and inspects equipment safely.

However, navigating unknown and dangerous terrains using a mobile robot with an egocentric teleoperation interface is difficult because the operator must build mental mapping to project themselves in the environment (i.e., difficult to achieve high situation awareness, Nielsen et al., 2007). The robots provide extended visual coverages of the distant environment using multiple specialized cameras for operators to increase their performance at tasks including search, inspection, and navigation (Kelly et al., 2010; Saakes et al., 2013).

Multiple cameras provide more remote information, helping an operator increase their situation awareness (Hughes & Lewis, 2004). However, while multiple cameras on a robot in teleoperation can contribute toward an operator's situation awareness, they may negate the benefit due to the increased cognitive load of the operator in comprehending the information (Drury et al., 2003; Kelly et al., 2010; Yanco & Drury, 2004b). Can we utilize the multiple specialized cameras to increase an operator's situation awareness of the remote environment in search and rescue teleoperation scenarios while reducing the disadvantages of increasing the amount of remote data?

Many multi-camera teleoperation interfaces use simple views (Lewis, 2013). The toggle view approach allows the operator to switch between sources on demand (e.g., Figure 1-3 on p.11 and Figure 3-1). Alternatively, the tile view approach presents multiple cameras at once (e.g., Figure 1-4 on p.12 and Figure 3-2). An operator has full control with

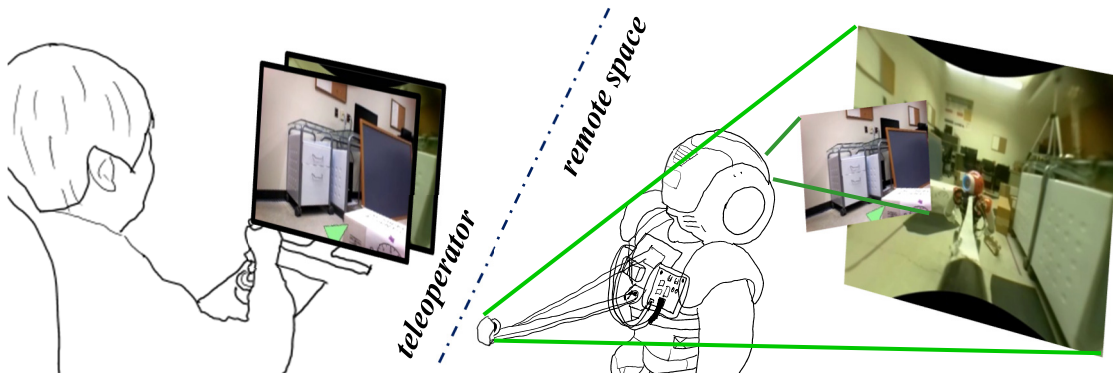


Figure 3-1. The mobile robot in the remote space (right) constantly provides sensor inputs to the teleoperation system (left) in real-time. With a toggle view interface, only one view is visible and other views are hidden. This figure conceptualizes the toggle view by putting one hidden view at other's back on the teleoperator's monitor.

toggle view interfaces and full access to every view simultaneously with tile view interfaces. However, with these types of teleoperation interfaces, the operator must build a mental map to understand the physical relationship between cameras and maintain awareness of other information (e.g., which camera is being displayed) to achieve their mission effectively. This extra cognitive load negatively impacts the operator's task performance (Wang Baldonado, Woodruff, & Kuchinsky, 2000).

There is a body of work for novel interface designs to provide more information while possibly not impacting a user's cognitive load significantly, including SLAM-based (Simultaneous Localization and Mapping based) or pre-built 3D maps (Labonte et al.,

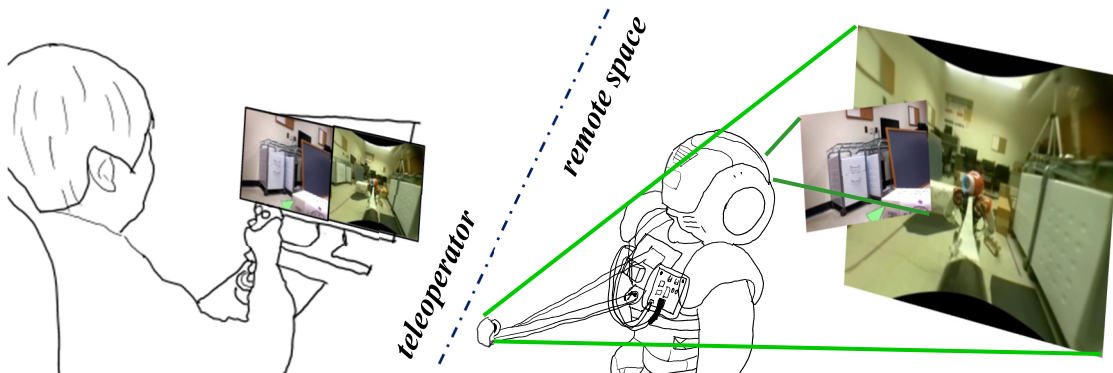


Figure 3-2. The mobile robot in the remote space (right) constantly provides sensor inputs to the teleoperation system (left) in real-time. With a tile (or side-by-side) view interface, on the teleoperator's monitor in this figure, all views are visible with reduced resolution on the screen.

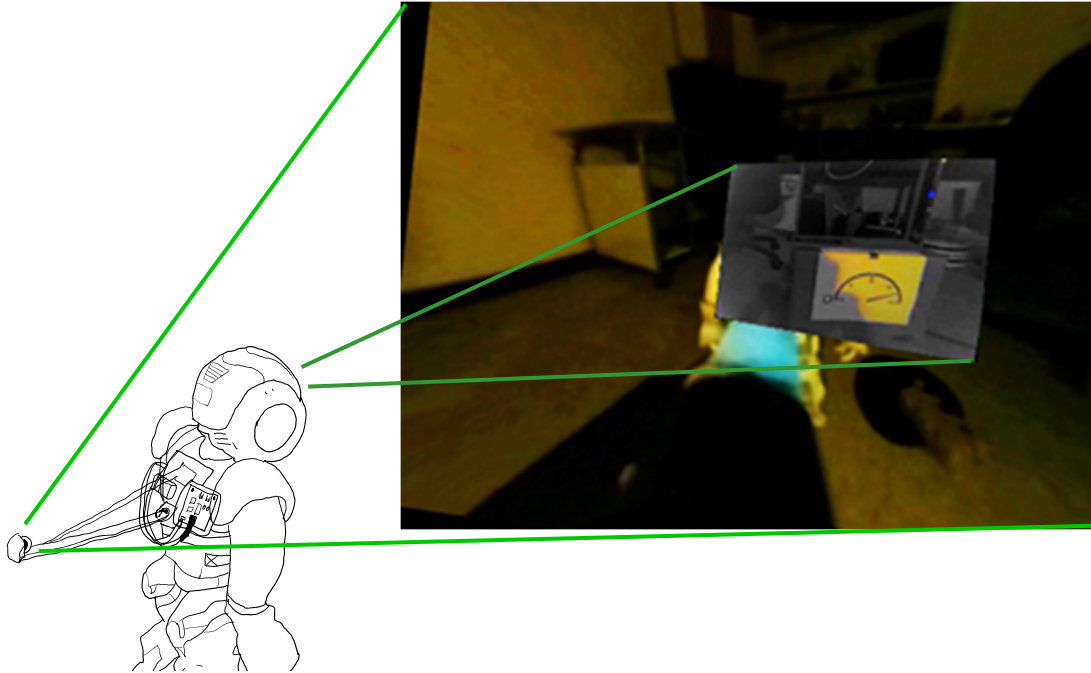


Figure 3-3. Monocle, our detail-in-context interface, a fully interactive detail-in-context teleoperation interface with wide pannable (behind the robot) and narrow pan-and-tilt (in the head) views. An operator can move the monocle around to gain detail where needed. (Low light, image quality due to search and rescue scenario).

2010; Ochiai et al., 2014), sonar (Salmanipour & Sirouspour, 2013), 3D generated contextual view (Richer & Drury, 2006), or rendered transformations and mark-ups in a virtual world to indicate the physical relationship between camera views (Ribeiro et al., 2016). This research creates and tests a novel interface for a remote mobile robot with multiple cameras. Our fundamental goal is to increase the operator's situation awareness and reduce cognitive load during search and rescue operations.

We leverage the information visualization paradigm detail-in-context (Keahey, 1998), also known as focus-plus-context (Baudisch, Good, Bellotti, & Schraedley, 2002), to have a single view with the appropriate detail interactively embedded in the context. The detail view overlays the environment view at the corresponding location, forming an integrated mixed-resolution feed, removing operator's requirement to map image feeds mentally (Figure 3-3 and Figure 3-4). We called our interface Monocle, because the detail view

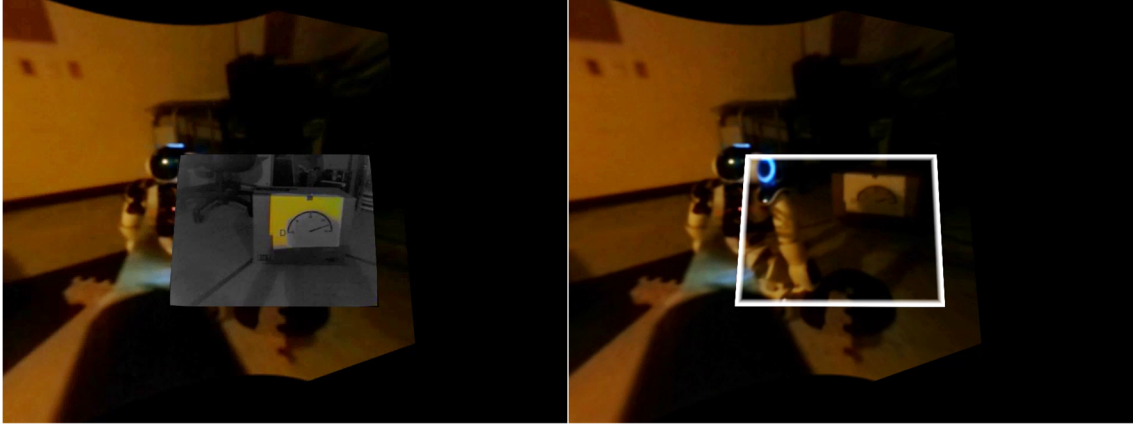


Figure 3-4. Monocle detail-in-context interface. (left) the sharp view integrates into the wide view; (right) the monocle can be toggled off by the operator. The third-person view provides a broad context, while the first-person view has a narrow field-of-view with more clarity.

acts like a monocle on a tabletop map that corresponds to our interface’s contextual view.

To test our novel teleoperation interface, Monocle, we use a mobile robot with two specialized cameras. We conduct a formal user study, compare Monocle to simple teleoperation interfaces, and list how Monocle helps an operator in a mock search and rescue scenario. We take light weighted (no SLAM processing or additional sensors), high fidelity (real-time camera feed instead of constructed environment), standalone (no environmental camera), and portable (no map required a priori) approach to implement Monocle.

Much of detail-in-context work deals with how to transition between scales visually. For example, between a large-scale map and a close-up region, a challenge is to use distortions that highlight the transition legibly (Böttger, Preiser, Balzer, & Deussen, 2008). However, our novel interface, Monocle, is a single scale but multi-resolution interface which bypasses the requirement of stitching mixed-resolution data sources like other work with mixed-resolution data sources (Baudisch et al., 2002).

While image stitching techniques similarly fuse multiple camera images into one (Kruckel et al., 2015; L. Lee, Romano, & Stein, 2000; Uyttendaele et al., 2004), our work has significant differences: (1) Monocle uses a mixed-resolution view instead of aiming

for one homogenous result image to maintain important detail, (2) Monocle is interactive in real-time with both views being user direct-able, and (3) we provide the results from a formal study of the use of this technique in a search and rescue scenario.

Monocle is the first system to integrate and evaluate two direct-able live camera feeds into a single mixed-resolution, real-time interactive display, providing detail-in-context interface that enables operators to simultaneously view the context of their robot operation while gaining fine details where needed, as needed. This solution helps an operator understand multiple camera views while not increasing their cognitive load significantly.

3.1. Monocle, the Detail-in-context Interactive Interface

We introduce our detail-in-context interactive teleoperation interface, Monocle (Figure 3-3). Monocle combines video feeds from multiple cameras into a single, calibrated view for the user. We used a robot with two cameras for our work, which provide a sharp narrow view (first-person, in-robot head) and a broad environment view (behind the robot). A robot with two camera views provides different details to an operator and helps them accomplish their tasks (Saakes et al., 2013; Seo, Rea, et al., 2017).

Narrow, first-person, and high-resolution cameras are useful for close inspection of equipment or potential survivors (Kruckel et al., 2015; Saakes et al., 2013; Singh et al., 2013). Broad environment views from the back of the robot show large areas, including the robot within the environment, helping an operator avoid holes and obstacles and visually search objectives (Hashimoto et al., 2011; Saakes et al., 2013). Such views can be achievable using a flying robot (Saakes et al., 2013) or a contextual view using an installation (Sakamoto et al., 2009). In this work, we did not rely on other video sources. We

implemented a standalone platform for the Monocle interface.

The multi-camera system becomes even more potent if an operator can steer each camera separately to look and get detail when and where needed. With Monocle, an operator can steer the high-resolution view to perform a visual search or equipment inspection, while steering the broad environmental view to plan their next move. The wide camera can pan, while the sharp camera can pan and tilt. The integration automatically updates as the cameras move, with the sharp overlay calibrated within the wide view. The operator can toggle on the sharp view to see through the robot or toggle off to see the robot within the environment (Figure 3-4).

3.2. Implementation

We use the Aldebaran NAO H25 humanoid robot, controlled using the NaoQi API and in-house remote-control software. While not a search and robot, it provides stable locomotion and camera operation appropriate for the implementation and evaluation.

We use the NAO's built-in first-person camera for our sharp view (pan and tilt control move the robot's head) and Raspberry Pi's camera for the wide view. We mount a light-weight boom (60 cm) and servo (120° pan) on the NAO's back to hold a 180° fish-eye lens (Figure 3-5); the boom points slightly downward to show the robot in the environment (Figure 3-6). We use OpenCV 2.4.1 to transform and merge two camera feeds.

Common techniques for stitching multiple images together (Zhang & Liu, 2014) use feature matching and transforms one image onto another. More recently, parallax-tolerant methods (Zhang & Liu, 2014) address cameras being at different locations, as in our case. We found current packaged image stitching methods unstable and not robust enough



Figure 3-5. Raspberry Pi's camera with no additional filter (left) and 180° fish-eye lens (right). With a fish-eye lens, the robot can capture more remote information; however, the details are missing since the camera's sensor and resolution are the same.

for use in our highly dynamic real-time streaming scenario with unstable lighting, a moving robot, and two steerable cameras through extensive testing. Even temporary or sporadic loss of calibration is unacceptable in search and rescue and would hinder our evaluation. We also need an agile algorithm that can be updated quickly to test our detail-in-context interface design idea.

Therefore, we implemented a simple stitching algorithm. The algorithm uses linear interpolation between calibrated transforms for each camera in the scene. We prepared and calibrated transform matrices manually for a set of camera pan and tilt combinations (e.g., head center and boom center, head top-left and boom left) to produce the most accurate alignment for each camera setting in human perspectives. During run-time, the algorithm continually interpolates and updates the matrices using constants such as the camera field of views and the physical configurations and run-time variables such as the camera odometry. We leave advanced image processing stitching techniques as future work.

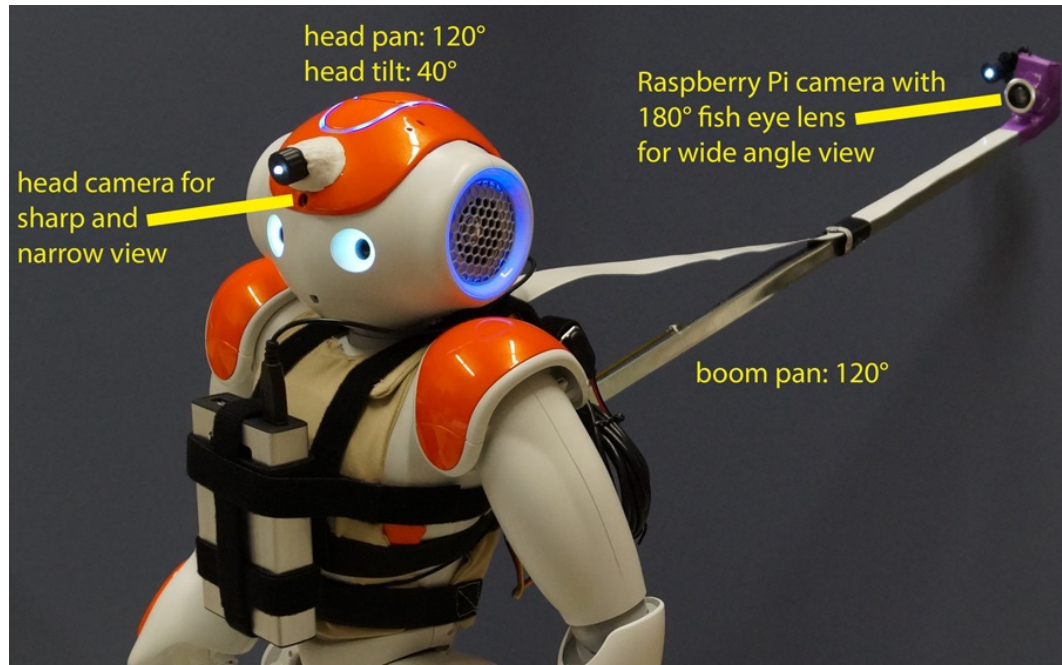


Figure 3-6. The robot with the boom; the boom camera is above the robot and points downward to capture the full body. Both cameras have LED lights attached to aid in dark environments.

3.3. Evaluation: Formal User Study

We conducted a formal user study to explore the advantages and disadvantages of our interactive detail-in-context interface. To confirm if our solution supports an operator's task performance while minimally impacting their cognitive load, the study compares Monocle to two widely used solutions: toggle view interface and side-by-side (tile) view interface. Toggle view interface provides one full-screen view at a time selectable on-demand by the operator. The tile view interface provides two views simultaneously. We administered various questionnaires (such as NASA Task Load Index) and measured people's teleoperation task performance in the study.

3.3.1. Comparison Interfaces: Toggle and Side-by-side Views

The toggle interface provides one full-screen view at a time (one of each in Figure 3-7) selectable by the operator. As the toggle interface utilizes maximum screen estate for a single view, it can be beneficial if an operator needs to focus on inspecting equipment or searching for victims using one camera feed. However, the toggle interface only shows one camera feed at a time, meaning that the operator does not receive as much information as the side-by-side interface provides at once.

The side-by-side interface (based on Singh et al., 2013, two views side by side, sharing screen real estate) provides two views simultaneously (as in Figure 3-7). As the side-by-side interface shows all views at once, the operator has the highest continual volume of remote information and does not need to track which camera is currently on the screen. However, to provide all views at once, the interface must reduce each video feed's screen real estate (i.e., the video feed's visual size on the screen, not the source resolution), which may not be desirable during visual search tasks.

We incorporated a compass-like feature to the toggle and side-by-side interfaces to indicate the robot's walking direction to the operator (bottom of the left image, Figure 3-7). Our pilot studies indicated that this information was critical for operation, as the robot's gaze direction does not match the robot's body direction. However, this information was not necessary within the Monocle interface because the sharp view (or white box if toggled off) shows where the robot is looking. The wide view shows the robot's environment and the robot in the environment.



Figure 3-7. Sharp narrow view (left), wide-angle view (right). The green arrow indicates the robot's facing direction (in contrast to viewing direction), blue border at the edge of the screen appears to indicate camera movement limitations. Both cameras' feeds are shown together in the side-by-side (tile) interface, but individually for the toggle interface. Note the victim, with simulated red blood, is easier to see in the sharp view.

3.3.2. Task

In search and rescue teleoperation scenarios, an operator must maintain a high level of situation awareness to complete their mission goals successfully, including navigating the area safely, finding victims, and inspecting sensitive equipment. Even with high situation awareness, careful navigation, and extra caution, testing an interface in a real disaster environment can damage robots severely and hinder our investigation of our interface. Therefore, we implemented a simulated disaster environment with fake holes, debris, survivors, and equipment gauges, focusing on our investigation.

In our simulated disaster environment (Figure 3-8), visual cues represent scenario instances. Black cutouts on the floor represent holes, cardboard blocks are debris that can trip the robot, plush toys with red tape (indicating blood) are survivors, and printed-out gauges on boxes are equipment items to be inspected. All boxes had a paper attached, and inspection was required to determine if a gauge is present. The room had low ambient lighting to improve realism, and the robot had two LED lights next to the cameras.

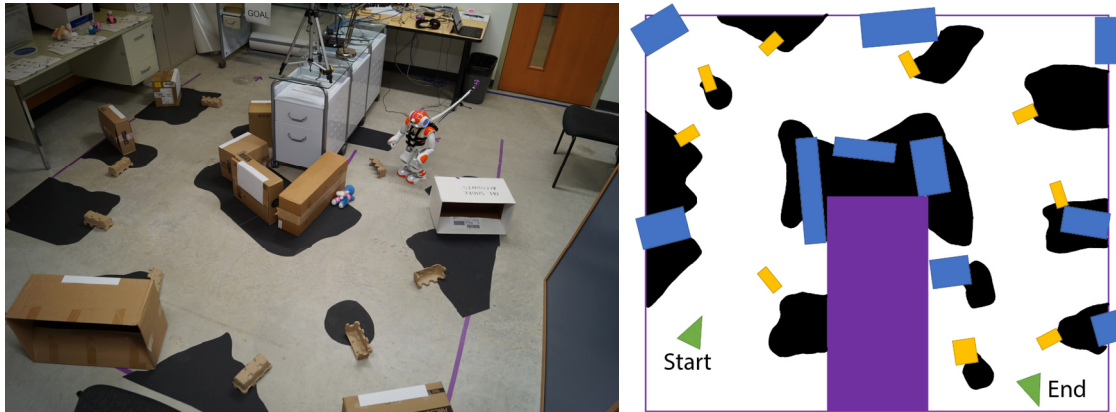


Figure 3-8. The picture (left) and map (right) of our search and rescue evaluation environment. The robot is at the start position in the picture. On the map, blue boxes are equipment, yellows are rubbles, the black surfaces are (fake) holes to avoid stepping on, and the purple box is a wall.

Participants had to navigate the environment from start to finish (Figure 3-8) and find as many gauges and victims as possible. While the route was reasonably linear, participants had to wander to inspect boxes, look behind obstacles, and avoid debris and holes. When the robot hit an obstacle or hole, the screen would pulse red to inform them of the problem. This feedback was necessary as the pilot study results suggested that participants would often ignore obstacles and holes, tripping the robot. For ease of implementation an on-site researcher monitoring the interaction provided this feedback manually, unbeknownst to the participant.

Participants need to record the reading and label of discovered gauges on a paper, which requires them to get close, and must also inform the researcher when and where they find a victim in the environment. The robot does not have to physically move obstacles or debris to find these, although some are more hidden than others.

3.3.3. Measures

This study investigates the impact of teleoperation interfaces on an operator's overall task

performance in search and rescue teleoperation scenarios. Teleoperation interfaces can be tuned for a specific task while removing other supports. For example, an interface can provide thermographic camera views from the remote environment to find victims quickly while not necessarily providing any navigation supports. Therefore, instead of measuring a specific task's performance (e.g., avoiding obstacles) in isolation, measuring overall task performance using a metric that combines the operator's performance of every task is essential to reflect the impact of the teleoperation interfaces.

Our primary measure is an aggregated performance score that considers the various components of the task, including how well participants completed the goals balanced with errors they made. We add up positive outcomes (e.g., finding victims) and subtract negative consequences (e.g., misreports or hitting obstacles). While we provide the detailed breakdown in our results for post-hoc discussion, we emphasize the importance of taking a more holistic view of operator performance. Our aggregated score begins with the sum of *the numbers of correctly reported victims* and *the numbers of correctly reported gauges*. From that sum, we then subtract *the numbers of accidents*, *the counts for misreported victims or gauges*, and *the numbers of missed victims and gauges*.

Accidents included colliding with an obstacle or stepping into a hole, misreports were either falsely identified victims or incorrect gauge readings, and missed items included gauges and victims.

In piloting the robot during the study, operator strategies varied between participants. For example, some lingered around or went back to the start location to check if they missed any victim or equipment, while others aimed to finish quickly. Given the task's

open-ended nature (i.e., we did not mention a precise number of victims or gauges to participants), we did not measure task completion time.

In addition to the performance score, we administered the NASA Task Load Index (TLX, Hart & Staveland, 1988) per interface, with additional questions about the perception of efficiency, awareness, nausea, and enjoyment.

Post-study, participants completed a written questionnaire that asks participants to rank the three interfaces (toggle, side-by-side, and Monocle) in terms of their preference and provide positive and negative thoughts for each interface and the user study. We also asked participants to verbally answer questions in a semi-structured interview, primarily asking about their previous teleoperation experience and their thoughts on the Monocle interface in search and rescue scenarios.

We extract people's general thoughts from the written comments and answers in the semi-structured interview and report emerging themes and ideas to provide deeper insights into the teleoperation interfaces.

3.3.4. Procedure

The joint-faculty research ethics board at the University of Manitoba approved this study. Participants completed an informed consent form and demographics questionnaire at the beginning of the user study.

Participants completed the primary task three times with each interface (toggle, side-by-side, and Monocle). There were three gauges and three victims placed in the environment; however, we did not mention the number to participants. The victims' and gauges'

locations changed between interface runs; the locations and the interface order were counterbalanced between participants. Before starting with each interface, participants completed a minor training on using the interface. Immediately after each interface, we administered the post-task questionnaire.

Each task ended when the participant reached the course end and announced they finished (they were allowed to go back), or a 12-minute hidden time limit elapsed. Participants were told that time was a factor, and the building may collapse, but we did not tell them how long they had.

After participants completed the tasks and answered the post-test questionnaire, we conducted a semi-structured interview to inquire about the overall experience. Finally, participants were debriefed on the experiment.

3.3.5. Apparatus

Participants used a PC with a 27-inch monitor in the lab environment and teleoperated a mobile robot in a different room (Figure 3-9, left). The chair and monitor locations were fixed at the beginning of the study; we allowed participants to move the chair or lean in or back to get comfortable.

To control the robot and the cameras, participants used a dual-axis Xbox 360 controller (Figure 3-9, middle). Participants used the left stick on the joystick to move the robot, the right stick to steer the sharp view (robot head), and buttons to pan the wide view (boom) left and right. To swap between sharp and wide view (toggle interface) or to toggle the sharp view on and off in Monocle, participants used shoulder buttons on the controller. A print-out legend of the controls (Figure 3-9, right) was provided to the participants.



Figure 3-9. A participant sits at the desk (left) and teleoperates a mobile robot located in a different room using a PC and a dual-axis Xbox 360 controller (middle). We printed out the control scheme (right) and placed it on the desk for participants to use during the study.

3.4. Results

We recruited 13 participants (3 female) from our university population and paid them \$10 for their one-hour participation.

Figure 3-10 shows the results of the aggregated performance scores by interface and the grand means with 95% confidence intervals across participants. We analyzed our results with repeated-measures one-way ANOVAs, with planned contrasts comparing Monocle to the two simple interfaces (toggle and side-by-side).

We found a medium-sized main effect of interface on performance score ($F_{2,24}=3.57$, $p<.05$, $\eta^2=0.15$) with (planned contrasts) Monocle ($M=-0.1$) scoring higher than toggle ($M=-3.6$, $p<.05$) and side-by-side ($M=-2.5$, $p<.05$). We found no interface effect on reported post-condition questionnaire data, NASA TLX reports, and no dominant preference for an interface.

As repeated-measures statistics rely on relative performance, we present the participant's relative performance in-depth with each interface in Figure 3-11. We sorted each participant's score per interface and counted how many people did best, second best and

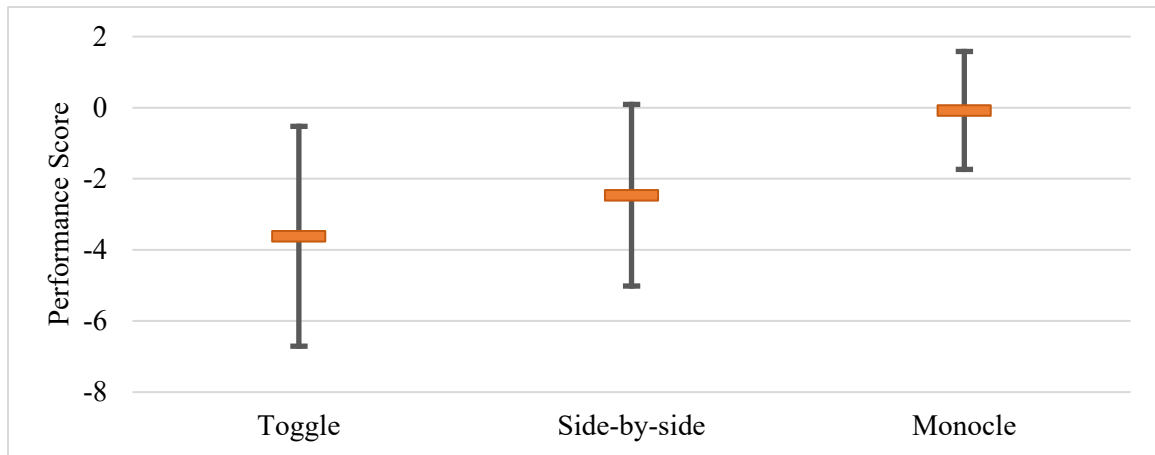


Figure 3-10. Overall average performance scores per interface with 95% confidence interval. The data skewed to negative values as there are many positive outcomes while infinite negative outcomes in our performance score aggregation.

worst for each interface. From the breakdown (Figure 3-11), we can see that more participants performed their best with the Monocle interface.

We conducted a qualitative analysis of the written responses and semi-structured interviews to understand each interface's advantages and disadvantages. We extracted dominant themes from the results using iterative open coding with initial codes inspired by our pilot studies. We used a single coder, given our exploratory and descriptive focus. As such, we did not conduct inter-coder reliability tests.

We first listed a set of initial codes to look for during the study from our findings in our pilot studies. Then, the coder iteratively categorized the researchers' notes written during the study and the points from participants' comments. Finally, we have collected dominant themes and reported them here with the number of participants for a specific theme and quotes.

Interface Simplicity and Usability

Participants commented heavily on the importance of simplicity of the interface, which

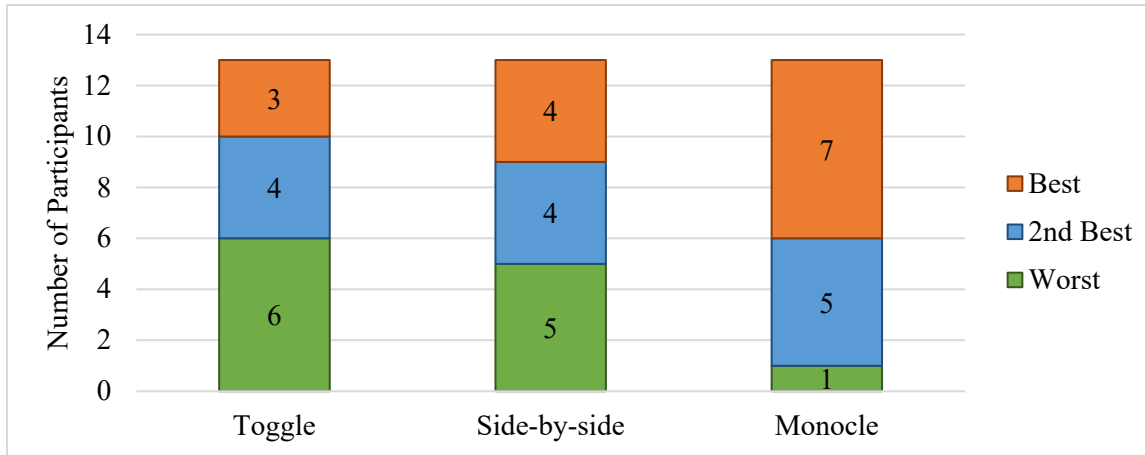


Figure 3-11. The graph illustrates how many participants did best, second best, or worst with each interface. For example, one participant performed worst with Monocle, while six participants performed worst with toggle.

was one of the primary benefits noted of the side-by-side interface (e.g., *“both views are available without switching”* 8 participants out of 13 participants). While participants also commented on the simplicity of the toggle interface (e.g., *“one thing to focus on at one time”* 7/13 participants), the requirement to toggle back and forth was seen as an issue that impacted performance (e.g., *“have to keep flipping back and forth”* and *“the toggle was not so convenient and time consuming”* 7/13 participants).

For Monocle, some disliked that sharp view occluded the view, even when it was only an outline (e.g., *“the front screen gets in the way”* and *“distracting view because of the outline of the merge camera”* 6/13 participants). A few participants felt that they would perform better with Monocle if they had time to practice (e.g., *“it takes a while to get used to overall [maneuverability] using the merged interface [Monocle]”* and *“felt easy to use after getting the hang of it”* 3/13 participants).

Supporting Awareness

Participants described how the interfaces supported their awareness of the environment,

such as the two views in the side-by-side interface (e.g., “*You can see more around you*” and “*can view the environment as first and third person at the same time, easier to avoid obstacles*” 9/13 participants). Users also felt this environment view to be an advantage of Monocle (e.g., “*still have third person view when using other camera makes looking closer at things easier*” 5/13 participants), which included being able to see through the robot via the lens (e.g., “*you can see in front of the robot and know where it is in relation to the back* → *easier maneuvering*”).

Participants also noted the awareness gains of toggle’s full screen view (e.g., “*larger visible area – easier to see,*” “*it’s good to have a bigger picture when inspecting,*” and “*it was more comfortable to see the view*” 8/13 participants).

Hindering Awareness

There were noted awareness-related issues with each interface. For toggle, participants noted that being unable to look at both views at once was a disadvantage (e.g., “*It’s could be ignored one side of environment [sic.]*” and “*easier to trip on something if you aren’t careful to go back-and-forth*” 7/13 participants). Similarly, while both views were available in the side-by-side interface, participants either found it distracting to have the two (e.g., “*Deciding which screen to look at was slightly more mentally tasking*” 7/13 participants), or that they would end up focusing only on one screen only (e.g., “*concentrate too much as one view at a time and forget about the other*” 7/13 participants).

For Monocle, some participants complained of the small size of the sharp view (e.g., “*much smaller if you want to look at detail*” 4/13 participants), which even impacted the perception of the robot’s capabilities (e.g., “*felt like the front screen was smaller [than other interfaces] in terms of motion and viewing side to side*”).

3.5. Discussion

Our results highlight that Monocle improved overall USAR task performance compared to more straightforward interface solutions. While the results were inconclusive on the NASA TLX and self-reported measures, this indicates that the detail-in-context view helped our participants complete their tasks effectively.

To gain a better sense of what impacted the overall score difference between the interfaces, Figure 3-12 shows the breakdown of the average scores. The number of accidents appears to be the most considerable change between interfaces. Intuitively, this matches the design goals of Monocle in providing clear environmental context while monitoring detail, without mental mapping, but further study is required to confirm this.

As important as the overall task effectiveness result, the participant feedback analysis provides insight into possible reasons behind each interface's performance. For a summary of these tradeoffs, see Table 3-1.

Feedback on Monocle highlighted that, even with a 27-inch monitor and close seating, the sharp view's smaller screen size may be a problem for operators. Due to limitations in network bandwidth and robot capability, our source image was 320x240 pixels. For all cases, the image is scaled up to match the overlap to keep all image resolution. Despite this, participants felt that they might see more if the overlay were larger. This remains an essential point for future work.

Although our results indicate that Monocle improved overall operator performance, it did not have a lower cognitive load (as per NASA TLX), and no preference emerged in participant feedback. However, as some participants mentioned, the interface seems unique

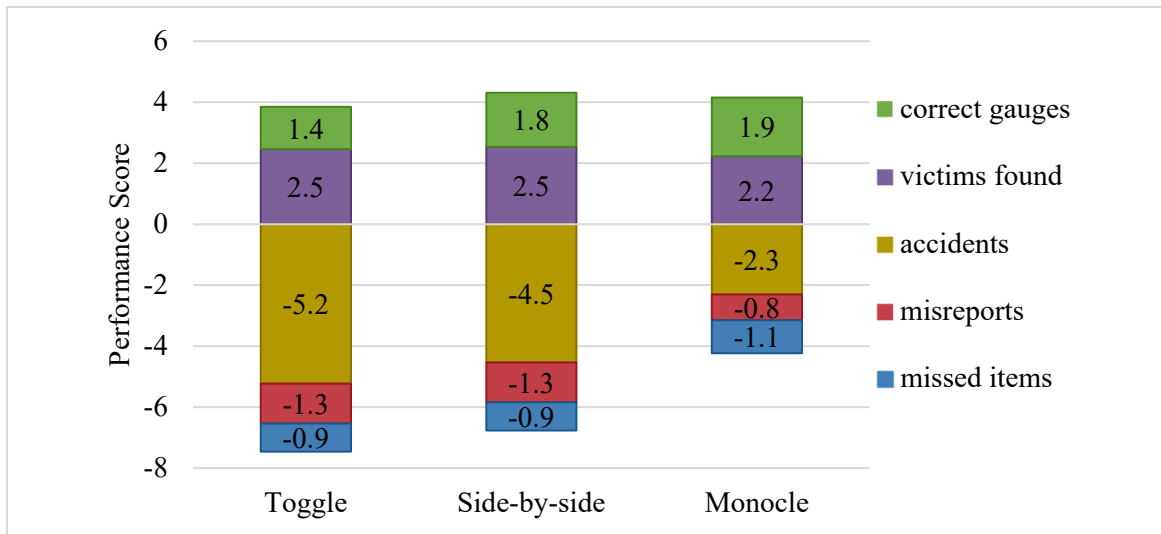


Figure 3-12. Average performance scores with breakdown of positive and negative factors. The number of accidents was the largest difference between interfaces.

compared to the other two. Monocle may require more longitudinal study with more operator experience before realizing all possible gains.

3.6. Future Work

Our method for combining two camera feeds using detail-in-context is just a starting point, and there remain many open questions.

Table 3-1. Summary of trade-offs emerging from the participant feedback analysis.

Interface	Advantages	Disadvantages
toggle	full-screen view for detailed search low effort to focus on one view	context switching (toggle to another view) less information (one view at a time)
side-by-side	two views support awareness simple to operate (no toggling)	two views can be distracting may over-focus on one view
Monocle	improved task effectiveness no mental mapping between views	small sharp narrow view occlusion of robot

Further improvements to the Monocle interface are needed to address raised concerns. For example, perhaps the small sharp-view could dynamically get larger as needed to enable an operator to examine something closely, perhaps using scale-transition techniques from detail-in-context. We also suggest that more investigation into modifications of Monocle that will mitigate the occlusion issue. One possibility is to provide a robot wireframe through the sharp view, or improved image stitching and alpha-blending merge.

Our interface only dealt with two cameras. Robots often have more, including cameras in robot arms, panoramic cameras, and so on. Extensions of Monocle to a generic space with more cameras is a non-trivial problem, particularly as cameras may be looking in entirely different directions.

3.7. Summary

Our primary goal is to assist an operator in achieving high situation awareness during search and rescue scenarios involving the remote operation of a robot having multiple cameras. We presented Monocle, a novel egocentric teleoperation video interface that leverages the detail-in-context paradigm to integrate two teleoperation cameras into a single mixed-resolution view, removing the need for an operator to map between the views mentally.

When we compared Monocle to the widely used teleoperation interface solutions (toggle and side-by-side views), we have shown that the detail-in-context technique improves improve an operator's overall task effectiveness during search and rescue scenarios. Further, our study results provide qualitative insight into future interface designs by comparing Monocle's advantages with other simple solutions.

Even with Monocle's success as described above, we have only touched the teleoperation interface challenges' surface including team awareness during multi-robot teleoperation. We introduce team awareness challenges, our novel interface designs, and evaluation results in the next chapter.

Chapter 4 Team Location: In-feed Embedded

Techniques for Visualizing Robot Team

Member Locations

We published a peer reviewed paper on this work and gave a public presentation at a conference. Some ideas and figures in this chapter are from that paper.

Stela H. Seo, James E. Young, and Pourang Irani. 2017. Where are the robots? In-feed embedded techniques for visualizing robot team member locations. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 522–527. doi.org/10.1109/RO-MAN.2017.8172352

Good teamwork involves proactive actions to help team members (Demir et al., 2017) and positive involvements between team members (Price & LaFiandra, 2017). Teleoperation robots, especially for search and rescue missions, are increasingly working in teams (J. Y. Chen & Barnes, 2013; Glas et al., 2012; Lewis, 2013) to increase search area, quickly respond to distress calls, and extract as much remote information as possible.

As in team teleoperation like many teamwork domains, good teamwork is accomplished through teleoperators maintaining awareness of team robot states, enabling team members' proactive assistance. Among many robot states, one state that can help operators maintain high team cohesion is the location information (e.g., Figure 4-1). For example, the operator can quickly approach a team member in need, or they can cover a wide search area by not overlapping their movement during search and rescue.

One widely used method for providing location information is a top-down map (Glas et al., 2012; Recchiuto et al., 2016; Zheng et al., 2013), often called mini-maps (Figure 4-2). Because of its wide use in many applications including teleoperation, navigation, and video games, we can expect that people are familiar with mini-maps and know how to use them when they appear in user interfaces.

However, mini-maps have limitations. The video feed has reduced screens space when placing the mini-map beside the video feed in egocentric teleoperation interfaces. If we place the map on top of the video feed, it covers the essential remote information (i.e., video feed). This is undesirable as the video feed is the primary source of critical information in teleoperation (Herring et al., 2016; Jaju et al., 2013). Whether the map placement is on a separate display or the video feed's display, the operator must perform a perspective shift between the video feed and the mini-map to understand both sets of information.

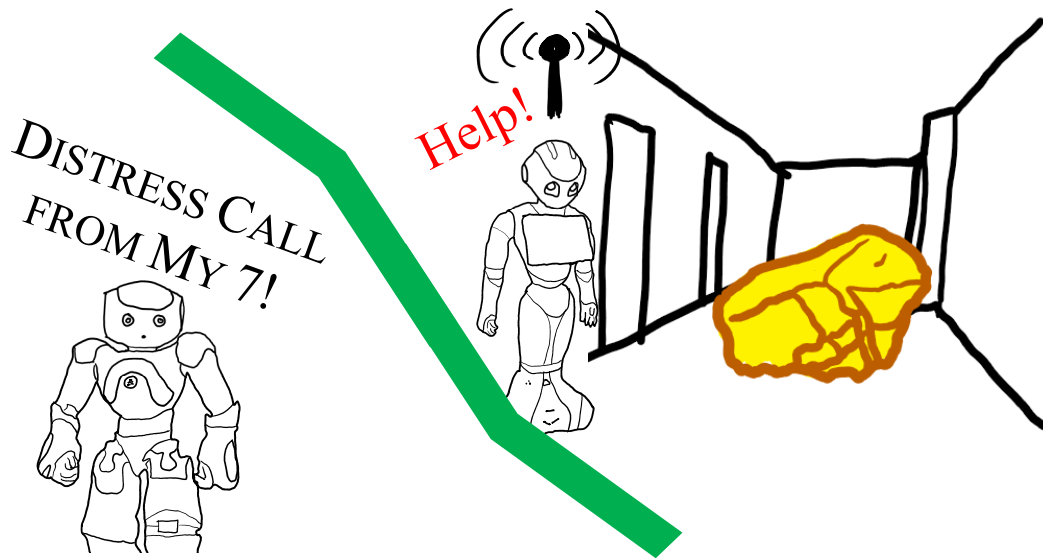


Figure 4-1. An example illustration when a team robot (right) finds a blockage and calls for help. If an operator, who controls a nearby (left) robot, is aware of the team member's location, they can quickly start moving toward the team member and save time to clear the path.

Our goal is to design novel techniques for conveying team robot locations while bypassing the limitations of mini-maps. That is, we aim to design location indicators that (a) reduce the operator's requirement to change their focus away from their task, and (b) reduce the impact on-screen real estate to maximize the camera feed's size and resolution, the primary source in egocentric teleoperation interfaces.

Our location indicators in egocentric teleoperation interfaces meet the above design goals by placing encoded indicators at the video feed's edge. This layout can help operators maintain awareness of their team-member locations using their peripheral awareness without requiring them to change their focus from the primary information source (i.e., the

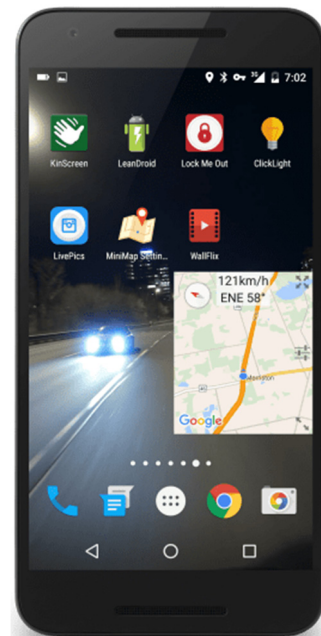


Figure 4-2. An example of mini-map on an Android phone. The mini-map takes a portion of the screen. Image retrieved from forum.xda-developers.com on August 7, 2020.

video feed). By minimizing the video feed's obstruction, the operator can also keep getting important information from the video feed. Since we have location indicators on top of the video feed, we call it in-feed location indicators of team robot members.

We introduce our design process and formal evaluation of our in-feed location indicators in an egocentric teleoperation interface in the following sections.

4.1. Designing In-feed Location Indicators

Our design goal for in-feed location indicators in egocentric teleoperation interfaces is to reduce an operator's context switching between the primary information source (i.e., video feed) and the map and minimize screen real estate taken by the location information. We believe that these two points are fulfillable with a carefully constructed design.

In searching for inspiration, we found a similar problem set in the field of mobile human-computer interaction. The field of mobile human-computer interaction has been working on embedded indicators for multiple off-screen items in an interface with a range of variants that target particular benefits and applications, due to the limited mobile display space (Baudisch & Rosenholtz, 2003; Burigat et al., 2006; Gustafson et al., 2008). Their off-screen target location indicators have the same design goal as our in-feed location indicators: maximize the primary visual content (in their case, it is the map content; our case, it is the remote video feed).

In our exploration, we have found three potential candidates that meet our requirements and appear to apply to the problem of both screen occupation and context switching. The three techniques are Halo (Baudisch & Rosenholtz, 2003), Wedge (Gustafson et al., 2008), and Arrows (Burigat et al., 2006). The general approach taken by these techniques



Figure 4-3. Two examples of off-screen target markers in video games. In the left example, off-screen targets (other players) are highlighted in red (*Overwatch*, Blizzard, 2016). The right example shows angled pointy markers at the edge of the mini-map with different colors for different types of off-screen targets (*Final Fantasy XIV*, Square Enix, 2013).

is to (1) encode both distance and direction of off-screen targets, (2) leverage the edge of the display, and (3) use people’s peripheral vision so that people can quickly compare two targets’ distance while being aware of their rough locational direction.

Another technique such as EdgeRadar (Gustafson & Irani, 2007) can be well-suited to moving robots. However, as with mini-maps, it consumes more screen real estate when compared with the three indicators. Furthermore, EdgeRadar assumes constant object speeds (Gustafson & Irani, 2007) unlike the examined scenario of dynamically moving robots. For these reasons, we excluded EdgeRadar from our implementation.

In addition to mobile human-computer interaction, we found that video games have extensive work on visualizing location information of multiple off-screen objects in a dynamic and noisy environment (Figure 4-3). While the lack of empirical research on video game methods makes it difficult to build directly on this work, we note that our selection of techniques from the mobile human-computer interaction literature with empirical results themselves build from video game design (Baudisch & Rosenholtz, 2003; Burigat et al., 2006; Gustafson et al., 2008).

Providing the locations of off-screen points of interest was also explored in driving navigation aids, primarily focusing on a single static destination with minimal divided attention, eye movement, and visual obstruction (Bark et al., 2014; Kim & Dey, 2009; Tonniss & Klinker, 2006). Like navigation assistants, the minimal obstruction can be an asset for location indicators to maximize the video feed in teleoperation interfaces. However, unlike an exact and static destination for a navigation task, we need to indicate the dynamically updating team robot location information. This focus on dynamics had led us to adapt three candidates for our in-feed location indicators: Halo, Wedge, and Arrow.

4.2. Our In-feed Team Robot Location Indicators

We adapt three location indicators, Halo, Wedge, and Arrow, to convey team robot members' location information. Specifically, we take team robots that are not visible in one's video feed as off-screen targets. We present those team robots' location information using in-feed location indicators (Figure 4-4).

Halo represents off-screen objects by drawing a circle around each off-screen item with a diameter calculated to barely overlap the intrusion on the view (Baudisch & Rosenholtz, 2003, Figure 4-5.A), such that halos are at the screen's edges. The direction to the object is simply the direction from the screen center to the halo edge, with the distance encoded in the curvature. Thus, the users can infer the circle center from the visible arc. A benefit is the encoded absolute location (distance and direction) in this fashion.

Wedge expands a triangle from each off-screen item so that the triangle base is visible on the edge of the viewport (Gustafson et al., 2008, Figure 4-5.B). Wedge uses the distance to the off-screen objects to calculate the base's length, such that objects further

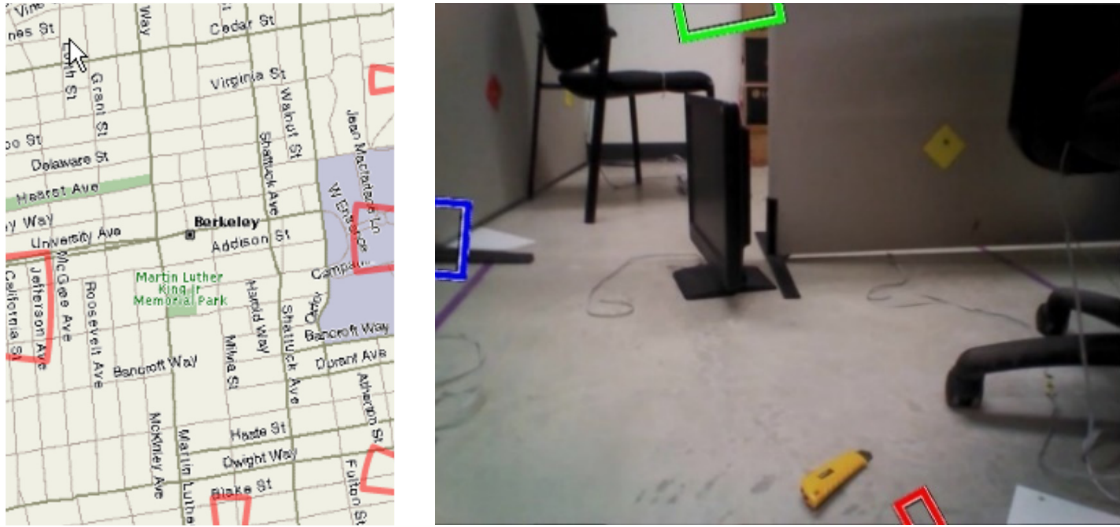


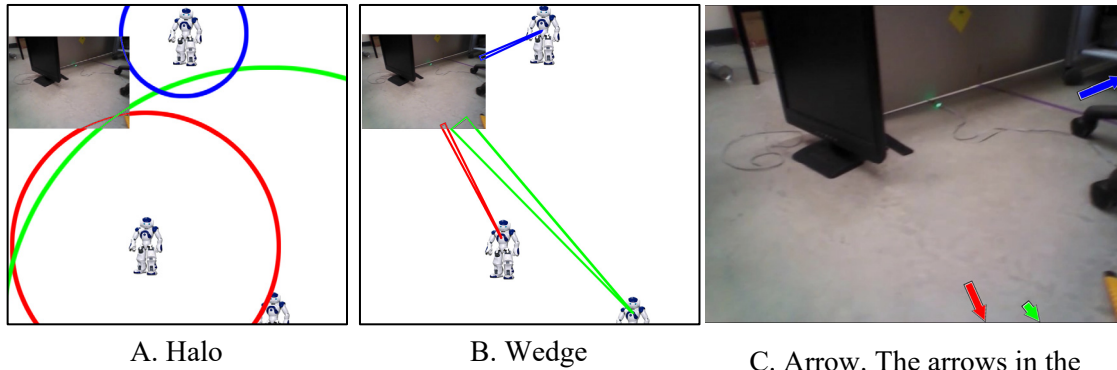
Figure 4-4. (left) Wedge (Gustafson et al., 2008) in mobile human-computer interaction literature shows off-screen objects. (right) We adapted and built an in-feed embedded indicator in first-person teleoperation view, showing team robot members' locations.

away have a larger base (and larger on-screen presence). As the triangle is from the off-screen object, the Wedge's location and the base's angle indicate to the user the direction toward the object. This information encodes the distance to the object as an observer can imagine where the triangle legs converge. Our primary goal in creating Wedge is to overcome Halo's clutter concerns (Gustafson et al., 2008).

Arrow (also called stretched-arrow, Burigat et al., 2006) places arrows on the edge of the viewport that point toward the off-screen object; the arrow is placed along the line from the screen center to the object, matching its direction. Arrow provides the object's distance using its tail length where the shorter tail represents the further object (Burigat et al., 2006, Figure 4-5.C). Arrow tail lengths are relative scaling linearly between pre-configured longest and shortest values, based on the robot distances.

4.3. Teleoperation Adaption Challenges

Adapting these indicators to the teleoperation task is non-trivial; there are challenges and



Both figures show the illustration of calculation.
The user only sees the video feed.

Figure 4-5. Our in-feed team robot location indicators on egocentric teleoperation interfaces. A and B show mock-ups of how Halo and Wedge work; only the viewport, with the indicators at the edges, are shown. C shows arrows that indicate closer objects with longer tails. In this scene, the blue robot is right and slightly in front of the robot, teleoperated by an operator.

essential differences that may hinder their performance.

All indicators assume that the screen represents a top-down view of an environment, as when using a map. This perspective directly conflicts with the first-person perspective commonly used in teleoperation. Embedding these top-down indicators into first-person video feeds requires operators to map between them mentally, which is not necessary for the original implementations. In our implementation, the up indicator represents forward, and the down indicator represents backward in the video feed (e.g., the green wedge in Figure 4-4 represents forward in the video feed). We note that this mapping is commonly used, for example, during the car navigation and video games.

Further, these indicators are primarily for mobile mapping applications having static maps and static points of interest. In contrast, robot teleoperation is dynamic, with both the primary robot (controlled or monitored by the operator) and team members are continuously moving. Some of which may move at varying speeds. In particular, when the primary robot turns, all indicators turn in the opposite direction to maintain their relative angles, creating a great deal of visual noise. This visual change results in constant indicator

movement, possibly causing distraction and reducing legibility.

These issues motivate the need to explicitly test the techniques in teleoperation scenario, and not merely rely on previous results.

4.4. Implementation

We implemented our in-feed embedded indicators with graphic overlays on a video feed, programmed using Unity3D. Given the video data's changing and diverse nature, we used red, green, and blue fills with white and black outlines to maximize indicator contrast.

One issue with our implementation was that the indicators' scale required changes compared to the original cases. The ratio between the viewport size and distance to robots was much larger in our case than in the mapping cases previously tested in mobile human-computer interaction. Previous work concludes that Halo does not scale well when objects are far away (Gustafson, 2008). The large distances caused all circles (even the closest one's) to become very large, appearing almost straight line. To solve this issue in our implementation, we scaled down the robot distances by 85%. Thus, unfortunately, our Halo loses its absolute robot distance benefit.

Like Halo, the triangle bases of Wedge implementation were often larger than the viewport and thus not useful when we were testing our implementation; we scaled down Wedge's distances by 75% to fix this, similarly losing its absolute distance benefit.

4.5. Evaluation: The Formal User Study

We investigate if our in-feed peripheral location indicators can help an operator keep being

aware of team robot member locations without hindering task performance in the multi-robot teleoperation context.

We conducted a within-participants study to explore how our in-feed indicators perform compared to the standard mini-map (Figure 4-6). In the study, participants completed a mock search and rescue task (the distractor) while maintaining their peripheral awareness of team robot locations and being regularly tested. The study included all four indicators: three in-feed location indicators and a mini-map as the base case. In addition to the four cases, we included a no-indicator case to examine how our indicators impact primary task performance (the distractor mock search and rescue task).

4.5.1. Apparatus

Participants used a PC with a 22-inch monitor, keyboard, and mouse. We fixed the initial chair, monitor, and keyboard locations, but participants were free to adjust them.

We used videos in our study to test people's peripheral awareness. The video provides consistency over time – every participant will have the same video feed regardless of their teleoperation skill levels. We were systematic in creating videos by entirely prerecording them in a mock search and rescue environment. In this fashion, we could ensure that only one light turns on at any given time on screen (green lights are for the light-searching task, see 4.5.2. Task). The videos were created by piloting an Aldebaran NAO H25 using a wireless connection and recording the teleoperation video, thereby ensuring the simulation's ecological validity. Five videos (with different environment configurations and light patterns) were prepared, with care to make them highly comparable, to enable five conditions in our within-participants study. Each video is exactly six minutes long.

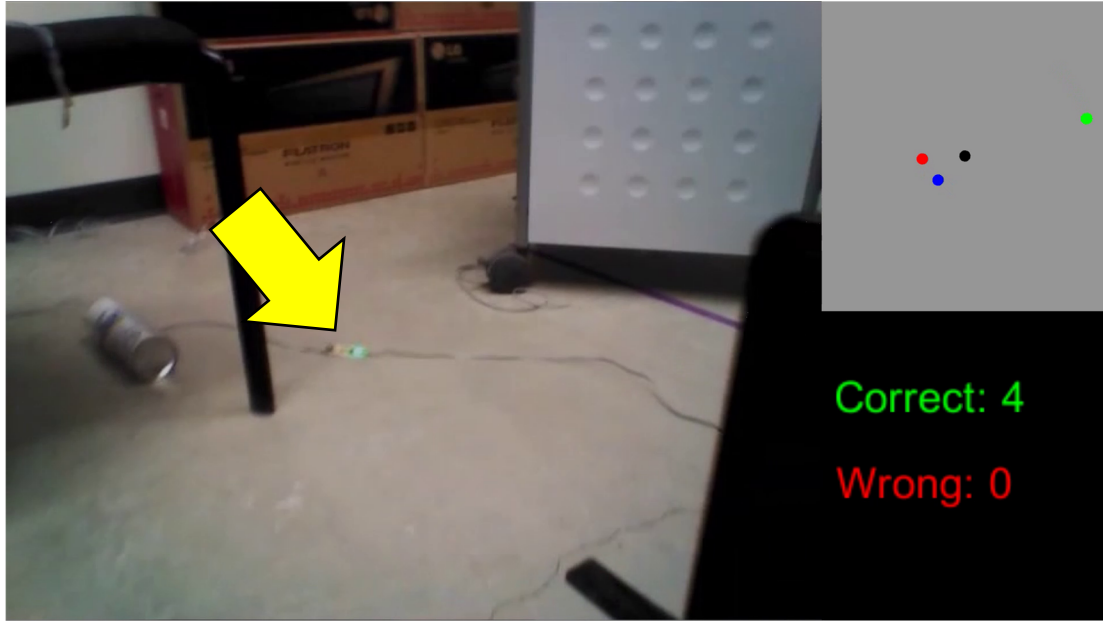


Figure 4-6. Interface used by participants. Participants perform the distractor task (searching for green lights, one shown, indicated by a yellow arrow in this figure) and hit the spacebar when one was found (score shown on the right), while maintaining awareness of the other robot members' location information. The mini-map was shown only in the mini-map case, with the region left black in the other cases.

The three team robots were entirely simulated (not told to participants), with the robot locations input into our indicators for visualization. The primary robot's movements were mapped and input into the simulation for movement and rotation reproduction. We also ensured that each video had a unique simulation dataset.

Our mini-map has three team robots positioned relative to the primary robot, represented by a black ellipse. This map is abstract, with a fixed scale (1:40) and no environmental details (Figure 4-6). The map automatically rotates to match the primary robot's gaze direction, to reduce the mental mapping. We placed the map next to the video feed to avoid occlusions and to provide as much video information as other cases.

4.5.2. Task

During the study, participants performed a mock search and rescue task where they were

simulating to control one of four robots. We informed participants that the videos were prerecorded, but we do not believe this hindered the results as they had a task to complete.

The task lasted for six minutes per condition, during which salient green lights appeared in the feed; these appeared on average every 30s but were unpredictable. We instructed participants were told that they had two tasks: searching for green lights and maintaining awareness of where the other robots are.

This distractor task, searching for lights, was used to raise participant cognitive load, divide their attention, and increase ecological validity. We instructed participants to press the spacebar when they found a light in the scene. To increase the distractor task's importance, if there indeed was a light and participants pressed the spacebar, a pleasant tone audio feedback played. The experiment interface played an unpleasant audio tone as feedback whenever participants missed a light or hit the spacebar when there was no light. At all times, the experiment interface kept their score on screen to encourage participants to focus on the light-searching task appropriately. The experiment interface (see Figure 4-6) only reveals the mini-map in the map case, and the region remains black during other conditions (Figure 4-7).

At unpredictable times, the interface paused the experiment, blanked the screen, and assessed participant awareness of the other robots. Instead of merely pausing the video, we blanked the screen to ensure that we assess peripheral awareness from an indicator and not overall legibility when given extra time to read it. The participants completed the assessments on-screen, instead of using a paper copy, to minimize the transition time from the screen to the questionnaire. The assessment happened three times per task, but we did not tell the participants how many times and when.

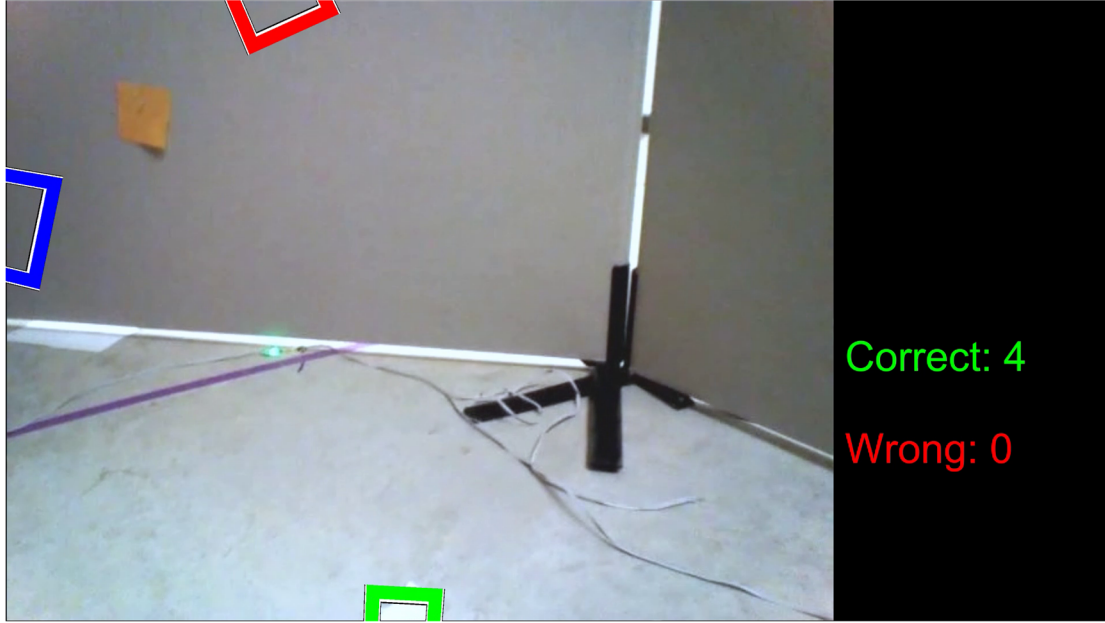


Figure 4-7. Screenshot of the Wedge condition interface used in the user study. In non-mini-map conditions, the mini-map (shown in Figure 4-6) was removed and the region filled black.

4.5.3. Measures

Our primary measure was participant awareness of the robot team members' relative locations. Given our task and interest in peripheral awareness, we are primarily interested in relative, coarse-grained peripheral awareness (which robot is closest to, which is in front) and not necessarily high precision information. We took this measurement using an on-screen interactive questionnaire, which asked participants to place the red, green, and blue robots at appropriate locations on a grid (for direction information). We then asked them to assess which robot was the closest and which was the furthest. We also assessed each participant's confidence in their answers.

We assessed the direction as correct if the response was reasonably close to the actual direction within the 11.25 degrees ($\pi/16$). This thresholding was a more appropriate measure of our interest in coarse-grained awareness and removed the unnecessary noise

of absolute measurement (e.g., 90 or 180 degrees off should be considered equally as wrong). A participant could achieve 0 for answering incorrectly for all three robots' directions to 3 for answering correctly for all three robots' relative directions.

Unlike the question assessing direction, the distance question asks participants to pick the farthest and the closest robots. As we measure their general awareness, we consider that people are aware of distance when they correctly determine which robot is closer or farther than another robot. In this experiment, we assessed the distance ordering question from 0 to 2 (i.e., 0=all incorrect, 1=one correct, 2=all correct for the closest and the farthest).

Pre-test, we administered a demographics questionnaire. Post-condition, we administered the NASA TLX to measure workload. We asked participants to report their level of nausea, sense of response speed for finding the lights, and how much they felt the interface demanded their attention, was distracting, or helped them maintain awareness. We asked participants to report the pros and cons of the interface.

Post-test (after all conditions), we asked participants to rank the indicators in order of their preference and provide any comments they may have relating to the indicators.

4.5.4. Procedure

Participants completed an informed consent form and were given honorariums at the beginning of the study. The researcher briefed the participant on the study and administered a simplified version of the Ishihara color perception test to exclude color blindness as a confound. After the pre-test questionnaire, the primary study started.

Participants completed the task with each of three indicators, the mini-map, and no indicator. While we fixed the order of the videos and simulations, the indicator orders were

counterbalanced using an incomplete Latin Square. For each task, the participants were trained to understand the encodings of the indicator.

The entire study took approximately one hour and a half and was approved by the joint-faculty research ethics board at the University of Manitoba.

4.6. Results

We recruited 21 people (6 female) from our general university population (mean age 26), who received \$20 CAD for their time.

4.6.1. Quantitative Results

We found no effect of having vs. not having indicator on the search task performance (light-finding accuracy, one-way ANOVA with planned contrast, $F_{1,20} = 2.1, p > .05$). The no-indicator case was excluded from all other statistical tests.

We use non-parametric statistics (Friedman's Test) on our ordinal awareness data. We found the effect of indicator type on awareness of relative robot distances is statistically significant ($\chi^2(3) = 11.2, p < .05$). The mean ranks were (lower scores are equivalent to lower performance): *Map*=2.83, *Halo*=1.81, *Wedge*=2.93, *Arrow*=2.43. Planned contrasts were conducted against the map base case: people recalled the relative robot distances more accurately with Map than Halo ($Z = -2.1, p < .05$). Post-hoc tests (Wilcoxon Ranks Test with Bonferroni correction) found that people reported distances more accurately with Wedge than Halo ($Z = -2.7, p < .05$). All other comparisons were not significant. Figure 4-8.A shows the within-participant performance of each indicator.

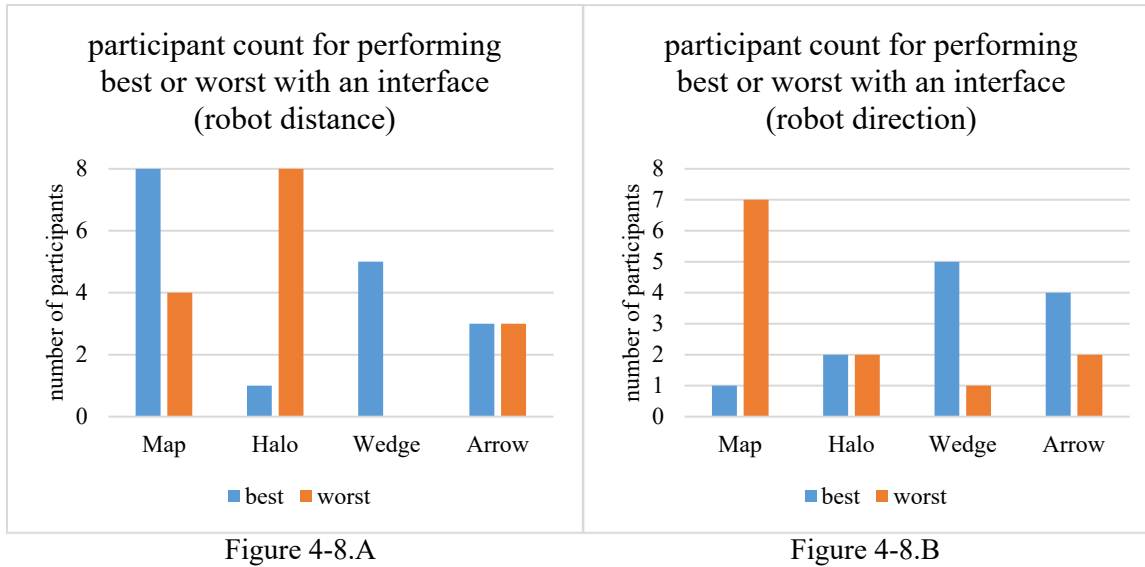


Figure 4-8. Charts represent relative performance showing how many participants performed, best or worst, with a specific interface.

We found a trend for indicators to impact participant awareness of robot direction ($\chi^2(3) = 6.4, p < .10$). The mean ranks were (lower scores are equivalent to lower performance): *Map*=1.95, *Halo*=2.69, *Wedge*=2.74, *Arrow*=2.62. Planned comparisons were conducted against the map base case, showing that participants recalled direction more accurately with Halo ($Z = -2.5, p < .05$) and Wedge ($Z = -1.9, p = .05$) than the map. Figure 4-8.B shows within-participant performance.

We found an effect on people's reports for each indicator type on how much it demanded their attention ($F_{3,60} = 2.7, p < .05$, Figure 4-9.A), with planned contrasts against the map base case showing that map ($Mean=8.2$) demanded more attention than Arrow ($Mean=4.9$). Other contrasts are not significant.

We found an effect for participant confidence in their distance reports ($F_{3,60} = 4.8, p < .05$, Figure 4-9.B), with planned contrasts against the map base case showing that participants were more confident with the map ($Mean=22.8$) than with Halo ($Mean = 15.5$). We found no other significant effects in this experiment.

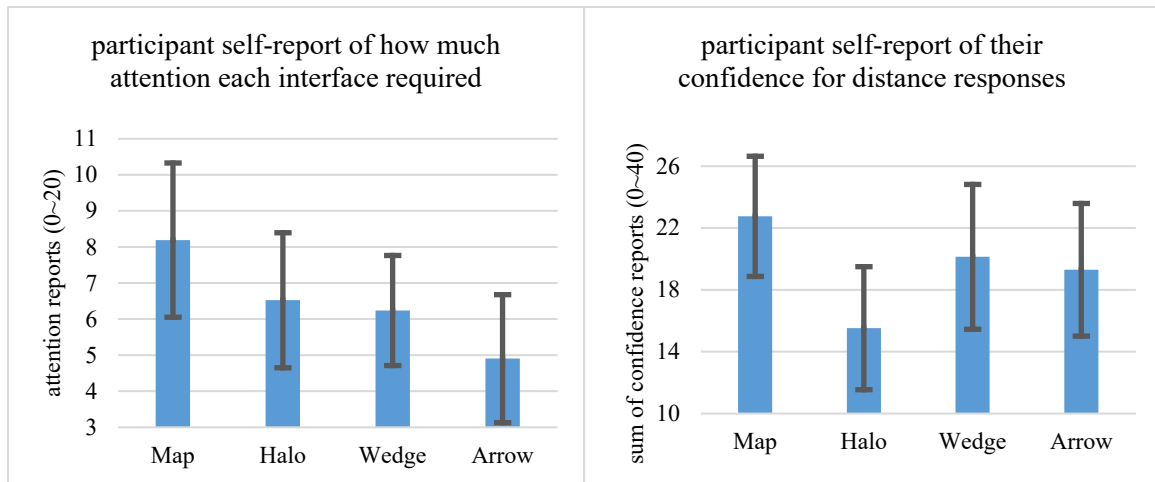


Figure 4-9.A

Figure 4-9.B

Figure 4-9. All scales were 0~20, with higher scores indicating more attention demanded and higher confidence. Confidence was measured twice (once for closest, once for furthest, robot) and summed. Error bars indicate 95% confidence intervals.

No effects were found on confidence in direction, cognitive load scores, nausea, sense of speed, or how much participants felt an interface helped them maintain awareness.

4.6.2. Qualitative Results

We conducted qualitative analysis on the written responses, from the post-task and post-study questionnaires to gain insight into some of the factors impacting indicator performance. We extracted dominant themes from the data using iterative open coding, with initial codes inspired by our pilots. We used a single coder given our exploratory and descriptive focus (no inter-coder reliability tests).

Readability

Many participants (15/21 people) explicitly mentioned how easy or difficult it was to extract information from an indicator. Nine found that Map was easy to understand for both the direction and distance (e.g., “*the scheme was easy to understand*” – P4).

In contrast, embedded indicators received fewer positive responses. Six people praised Halo for reading direction (e.g., “*easy to find the direction of other robots*” – P20), six praised Wedge (e.g., “*direction was easy to determine*” – P14), and eight praised Arrow (e.g., “*very easy to tell what direction the other robots are away from you*” – P9). Halo did not receive any praise for reading distance, whereas eight praised Wedge (e.g., “*easily able to tell which robot is the closest to my position*” – P10), and five praised Arrow (e.g., “*easier to tell the distance*” – P8).

Ten people said Halo was challenging to understand (e.g., “*not much cue for distance and specific direction*” – P15), four said Wedge (e.g., “*hard to tell the distance*” – P7), and four said Arrow (e.g., “*the size is difficult to see*” – P14).

In-feed Integration vs. Separation

Four people commented on the benefits of the integration of the embedded cues (e.g., “*the cue [Wedge] is part of the screen so you see them simultaneously with the green light*” – P17). Inversely, eight people commented on the problems of the separation with the map (e.g., “*in a way it [Map] was like a rear view mirror in a car ... it was so detached that it felt like I had to take my eyes off the road to make a note ...*” – P13).

Difficulty and Confusion

Ten participants noted difficulty distinguishing distances with Halo (e.g., “*very difficult to tell the direction because it is hard to distinguish curvature quickly*” – P14).

Further, two mentioned difficulties with cues (e.g., “*the wedges can overlap each other so it’s harder to distinguish*” – P3), and one noted confusion with the length (e.g., “*I intuitively thought longer arrows would mean further away*” – P3).

Attention Grabbing vs. Distracting

Only a few participants mentioned distraction. One mentioned Halo was not distracting (e.g., *“not distracting when it comes to telling where the robot is from your position”* – P10), whereas another participant mentioned it is distracting (e.g., *“very distracting”* – P20). Similar words were mentioned for Wedge (e.g., *“less distracting than other cues”* – P20), with two people finding it (e.g., *“kind of distracting from the cue”* – P9).

Two participants reported that the map grabs too much attention (e.g., *“not sure why but I find I was pretty absorbed into the cue [map indicator] that sometimes I missed the goal [light]”* – P3). No participants mentioned this for other indicators.

4.7. Discussion

Overall, our results indicate that embedded, in-feed indicators serve as viable alternatives to mini-maps for helping an operator maintain awareness of robot team members’ locations. All three in-feed embedded indicators performed as well as or better than the map alone for awareness of robot direction. While the map appears to be the strongest performer for awareness of distance, Wedge and Arrow performed well enough compared to the map.

Among the in-feed embedded variants, there are signs that Wedge may be the best performer. The raw score data suggests that Wedge may have performed better among all indicators, including mini-maps, although further experimentation is required. Halo was generally a poor performer, as participants commented on, although it still performed better than the map for direction awareness.

While overall no indicator outperformed the mini-map, both our quantitative results

and participant feedback suggest that the map may demand more attention and require more operator monitoring than the others. In contrast, participant performance did not reflect this, and participants reported reasonably high confidence in using the map. One possibility is that our primary task's difficulty was not sufficiently high enough to impact the participant. This can be a good variable for future formal investigations.

The indicators are aligned such that the 3D world's surface direction information is in a 2D top-down orientation (top as forward and down as backward). We believe that as many applications use this (e.g., navigation), people may be familiar with such translation. In future application work, we need to explore a useful perspective mapping for the embedded indicators to convey full 3D location information.

This work is a proof-of-concept step for using embedded indicators to support peripheral awareness of robot team members' locations. Further investigation into in-feed indicator designs for targeted tasks in remote operation is needed. Two key areas are to improve in-feed indicators' peripheral legibility and to utilize a 3D perspective to move beyond the top-down paradigm.

4.8. Summary

We demonstrated how in-feed embedded indicators of robotic team member locations could be applied when teleoperating a robot. Compared to a mini-map, which is a common solution for location information, in-feed location indicators use less screen estate, demand less attention, and require less mental mapping – our in-feed indicators occupy <1% of video estate while the map takes 18% if overlain. Our in-feed indicators also reduce the operator's context switch by leveraging their peripheral vision.

Our results provide insight into some of the trade-offs between the in-feed embedded indicators, such as being better for distance or direction information or requiring more or less operator attention. These outcomes provide a roadmap for future work, continuing to improve location indicator design using peripherals.

Overall, our work contributes to teleoperation interface design, providing a way to support operators as they operate an increasing number of robots in the remote environment. Specifically, our work shows one design to support the operator's situation awareness by providing the primary information source (video feed) and team-member location information with the minimal obstructive visualization.

Our in-feed indicators are a good alternative to the mini-maps, especially for egocentric teleoperation interfaces in team teleoperation, which use the video feed as the remote environment's primary information source. That is, even though the in-feed location indicators take much less screen estate, they still provide as much location awareness as the mini-maps do.

While our implementation and evaluation provided design suggestions of team robot location indicators to interface designers, we did not explore how other possibly overwhelming teleoperation information regarding team robots (e.g., team robot's robotic states and working states) can be presented in egocentric teleoperation interfaces. Thus, we continue our exploration in designing team robot state indicators.

Chapter 5 Team States: Graphical Techniques for Maintaining Awareness of Team Robots

We published a peer reviewed paper on this work in a journal. Some ideas and figures in this chapter are from the paper.

Stela H. Seo, James E. Young, and Pourang Irani. 2020. How are Your Robot Friends Doing? A Design Exploration of Graphical Techniques Supporting Awareness of Robot Team Members in Teleoperation. *International Journal of Social Robotics* (2020). doi.org/10.1007/s12369-020-00670-9

In team teleoperation, for productive and effective teamwork, an operator needs to maintain their awareness of team members, including team members' location, progress, states, notifications, errors, and so on, while pursuing their mission goals. In the previous chapter, we presented our in-feed indicators that provide team robot location information with minimal obstruction of the primary information source. However, with additional team information (e.g., team robots' internal states), the operators have a higher cognitive load to understand the information – this can increase the probability of human error and negatively impact their performance. Interface designers need to avoid this higher cognitive load because user error remains a primary cause of teleoperation mistakes (J. Y. C. Chen et al., 2007; Gombolay et al., 2017).

Our effort lies in increasing an operator's awareness of team robot members to avoid human mistakes when controlling a mobile robot using an egocentric teleoperation interface (Figure 5-1). We present prototypes that simplify complex robotic team robot states and actions, integrated into the interface, aiming to help the operator understand the team robots easily. Our prototypes consist of different information representations such as text, icon, and facial expression and graphical visualization parameters such as color, animation, and the number of team robots.

To test people's awareness of team robot states using our implementations, we conducted a series of pilot studies to explore a range of representations for these parameters (e.g., emotion, icon, use of color, motion, etc.). In these studies, we asked participants to maintain their awareness of team robot members while performing a task in a simulation through a mobile robot's remote operation interface. We iteratively explore different team robot state representations through a series of user studies.

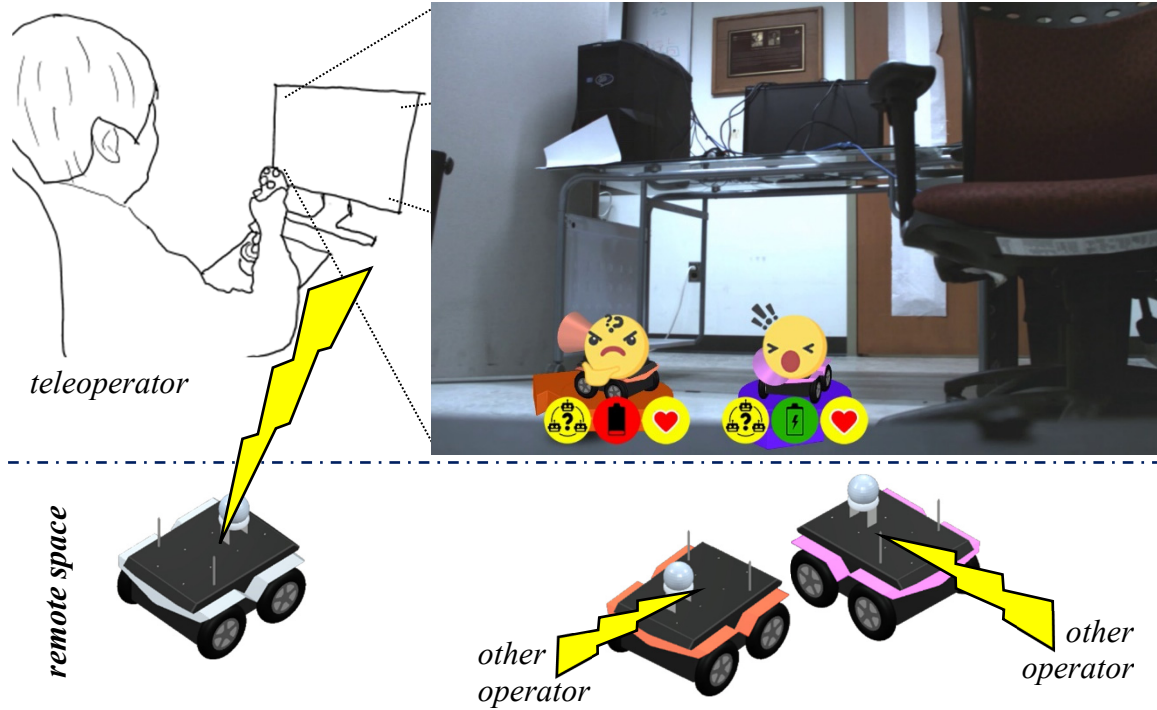


Figure 5-1. An operator teleoperates a robot located in remote environments, with two additional robot team members. While controlling one robot, they maintain awareness of other robots using on-screen widgets (screen bottom).

We conclude by listing a set of design guidelines from the results of our iterative exploration. We expect the guidelines to help designers of advanced widgets better represent team robot states in team teleoperation. We additionally discuss our findings regarding the advantages and disadvantages of our team robot state representations. Our discussions and design guidelines can help interface designers make informed decisions on designing team robot state representations in teleoperation interfaces.

5.1. Graphical Representations of Robotic States in Multi-Robot Team Teleoperation

We explore novel graphical widgets to support an operator to maintain awareness of other robots' states in their team while controlling one robot (Figure 5-1). We focus on egocentric

video teleoperation interfaces, where the display shows a camera feed from the primary controlled remote robot, and on-screen indicators and widgets presenting the additional pertinent information about other team robots.

To enable iterative exploration of widget representation for team robot states, we first devised a set of robot states representing information broadly pertinent to working with multiple robots. We keep this information generic to support generalization across teleoperation contexts. This constraint enables our results to apply to other teleoperation scenarios.

Following our discussion on generic robot states, we explain our general visual design approach and detail our specific strategies for encoding the given robot states into visual indicators in our interface.

5.1.1. Template Generic Team Robot States

Teleoperation is broad, vast, and task-dependent challenges. In various domains, from hobby activities (e.g., drone racing) to military activities, we see different challenges and problems. Instead of covering the entire field, we scale our work to search and rescue. An operator controls a robot in real-time to perform tasks while simultaneously monitoring activities and states of other robots in their team. We focus on general information that is likely to be important across a broad range of teleoperation tasks involving multiple robots in search and rescue scenarios. With this in mind, we settled on representing robotic movement and activity, abstract task information, and essential robot status (Table 5-1).

Robot Movement and Actions – it is essential to know other team members' general

movements and actions. We convey whether a robot is moving or not, its speed and direction. Further, we convey the robot camera’s gaze direction and movement, for example, when a robot is not moving but is carefully surveying a scene.

Abstract Task Information – to increase generalization of our exploration we avoid targeted task-specific information (e.g., finding a victim, space coverage, inspecting equipment) and instead abstract task information to a robot having a message to share. We convey the state of whether there is no message available, a message is available, or an urgent message available.

Robot State – we selected three generic internal robot states that are broadly relevant in teleoperation: network connectivity, battery level, and system failure (or physical damage). Connectivity is essential for expectations of responsiveness, and whether the shown information is current. Battery level is typical information for all mobile robots and broadly represents the ability to continue functioning. Finally, robots are fragile and have system errors or receive damage from the environment, explaining erratic or poor behavior.

Table 5-1. General robot states. We explore variations of the four robot states: connectivity, battery level, physical damage, and mission messages. We abstract the details into three categories.

<i>Robot Movement and Actions</i>	moving		not moving	
	looking around		not looking around	
<i>Abstract Task Information</i>	mission message	no message	message	urgent message
	connectivity	strong connectivity	okay connectivity	weak connectivity
<i>Robot State</i>	battery level	strong battery	okay battery	weak battery
	physical damage	no damage	light damage	heavy damage

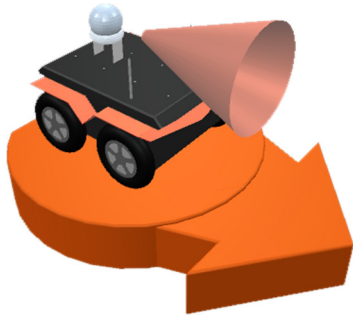


Figure 5-2. A team robot indicator with its moving direction (arrow on the bottom) and gaze direction (cone on the top).

While we accept the limitations of our generic robot state selection and note the importance of continued exploration into more task- and robot-specific states, these generic states are useful for examining indicators across a range of teleoperation tasks.

5.1.2. On-Screen Visual Representation Strategy

To develop our on-screen widgets, we worked closely with a local design firm, ZenFri Inc., and held a series of informal design sessions. We quickly converged on widget designs to represent the robot movement and actions during our design meetings, as this information is straightforward to show visually. As shown in Figure 5-2, we use a 3D model of our robot, placed on a large direction arrow to show its orientation. The wheels move prominently to show their movement, and the cone in front of the robot moves to indicate the gaze direction and camera movement. Given the team consensus on this indicator, we did not iterate the design further and instead focused on exploring the abstract task information and the robot states.

During our design sessions, we settled on three primary approaches for visual widgets representing team robot states that emphasize simplicity: short text messages, icon representations, or emotional information encodings (emojis). Further, we explored whether the designs should be in color (with meaningful encoding) or greyscale, should be animated (to draw attention) or static, and how many robot widgets should be on a screen at once.

Using these representations and graphical visualization parameters, the interface

overlays the team robot’s four states (Figure 5-3, Figure 5-4, and Figure 5-5).

Text-Based Robot State Representation

We selected text as a standard approach to conveying information, that is broadly understandable within a language group. The text further serves as a base case, an existing common approach, against which to compare our new methods. We would expect the text to be slower to understand than our other icon methods as it requires reading.

For simplicity reasons, we restricted each state to be represented by at most two words (one adjective and one noun, Table 5-1). We tilted the text 45 degrees (Figure 5-4) to maximize its physical size within a small screen footprint and match the other methods (square area). We added a black outline to maximize contrast and support readability. We choose text color representing standard cultural meaning (e.g., green is okay, red is danger; see section Color Coding of Robot States).

Icon-Based Robot State Representation

Icons can metaphorically convey a complicated message (Gittins, 1986) and can be quickly



Figure 5-3. Text representations for all states conveyed. Each state consists of two words: a noun and a descriptive adjective. Text is tilted to maximize size within a square space, the same size with icon representation (see Figure 5-4). We used a black outline to increase contrast and maximize visibility.

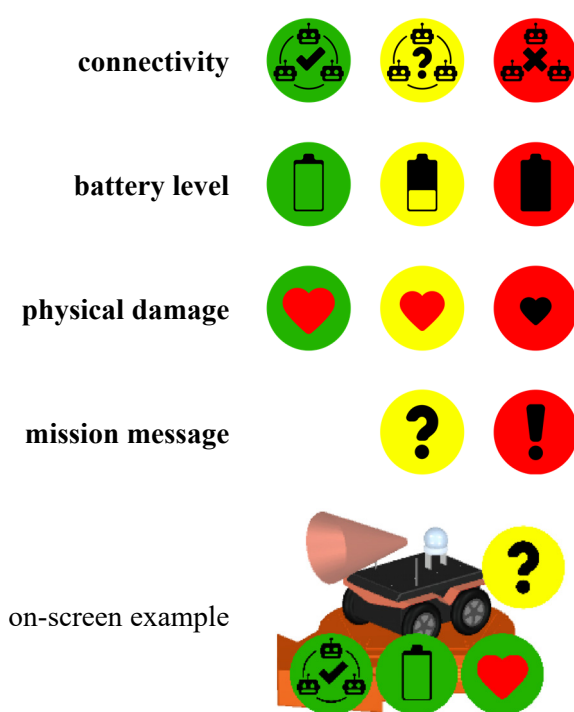


Figure 5-4. Icon representations have one-to-one mappings to the text (see Figure 5-3). The icons are placed around the robot model, in the same position as the text indicators.

understandable even when small or not in immediate attention (Kline et al., 1990) and while moving (Long & Kearns, 1996); well-designed icons can quickly and clearly convey meaning. In comparison to text, we expect icons to be quicker to understand and perhaps easier to interpret.

For connectivity, we focused on showing robots being connected to each other, battery level uses a familiar partially-filled battery icon, physical damage uses hearts as is common in

video games, and mission message uses a question mark to indicate the robot wanting the operator's attention (Figure 5-4). We chose icon color to represent standard cultural meaning (e.g., green is okay, red is danger; see section Color Coding of Robot States).

Emotional Information Encoding

We developed emojis that use facial expressions to represent the overall combined robot state generally. We envision that adding emojis can increase communication bandwidth because aggregated facial expression can be used to convey robot state and communicative signals through leveraging human social signal processing (Hess & Hareli, 2015; Parkinson, 2005; Sharma et al., 2013; Singh & Young, 2013). With a team robot's facial expressions, we expect people to take a quick glance at the representation, build a general idea of the

robot's current states, and further increase their overall awareness.

Our team robot facial expressions (i.e., emojis on team robots, Figure 5-5) aggregate all states into a single facial expression. That is, we assigned a numerical value for positive, neutral, and negative states, summed values, and grabbed assigned emotional encoding state based on a set of predefined ranges. For example, if all states are positive, the robot shows a happy emotion. We used two colors: red (angry) for overall negative states and yellow (all other emotions) for overall positive or neutral states. We did not include green because operators could interpret a green face to be a sick face.



Figure 5-5. Emoji shows the condition of general states using a facial expression. The message state is embedded in facial expressions. The bottom example shows that a team robot has all positive states and no messages.

5.1.3. Graphical Visualization Parameters

In parallel to our three design approaches (text, icon, emoji), we explore general graphical design parameters and their impact on the resulting widget: color coding (vs. greyscale), animation (vs. static), and several widgets on screen at once.

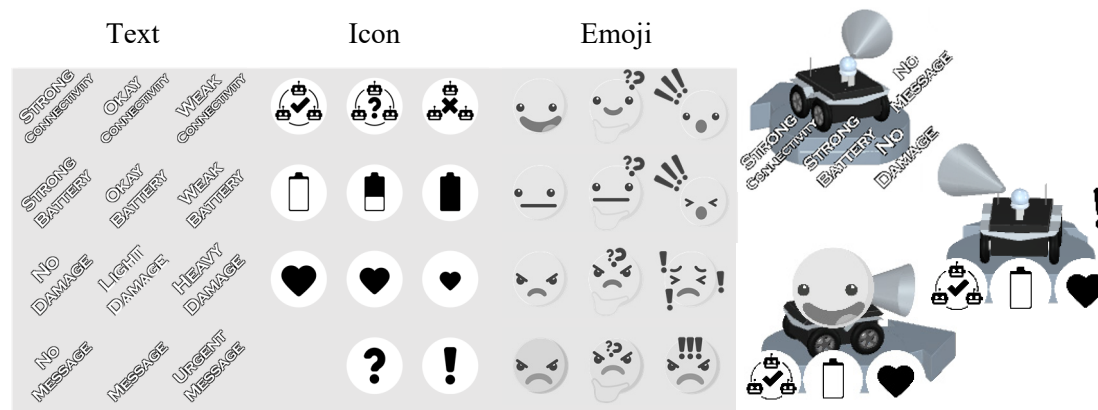


Figure 5-6. Every asset in the greyscale version: text, icon, and emoji representations. In this version, the team robot is in greyscale. Examples with each representation are on the right.

Color Coding of Robot States

Color can intrinsically carry meaning or be used to increase contrast and visibility (Murch, 1985). However, it adds visual complexity and may be distracting. As such, we explore the use of full-color techniques versus greyscale variants (Figure 5-6). In general, we mapped positive states to green, negative to red, and an intermediate to yellow to follow common cultural standards. For emojis only, we used two colors, red and yellow as we described in the above section (Emotional Information Encoding).

Animated and Non-animated Icons

The movement and change in animation can draw a person's attention, help them notice a state change or pay attention to a widget, but this can be distracting and negatively impact performance (Rea et al., 2017). We investigate both static (perhaps not distracting but might be ignored) and animated (might be distracting, but easy to notice) variants of our icon and emoji interface (Figure 5-7). Our animated case is continually moving within the fixed-size frame like a game agent.

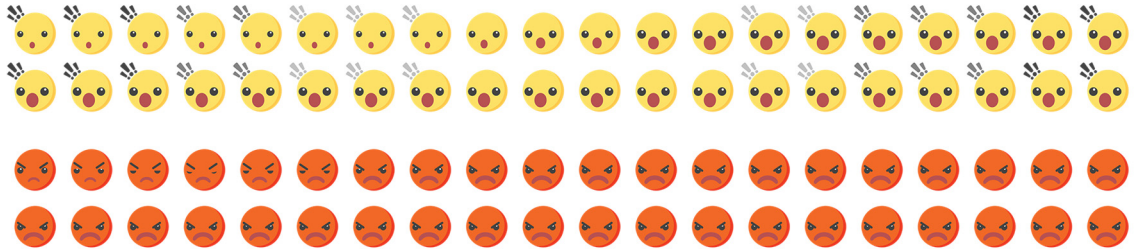


Figure 5-7. Two sprites (happy with an urgent message and angry with no messages) – Each sprite has 40 frames, which provide smooth animation. One cycle of the animation takes about 1.33 seconds. The animation continually plays in our animated representations.

Number of Team Robots

We anticipate that the operator’s cognitive load will naturally increase as we increase the number of robot widgets on the screen. To investigate this, we compare having one on-screen widget, representing having one team robot in addition to the operator’s primary robot, against two on-screen widgets, representing having two additional team members.

While the two widgets are functionally identical, they display different information as they are representing two different robot team members. We also changed the base color, with one robot being orange and the other being purple (Figure 5-8).

5.2. Exploration Testbed

We developed a novel testbed to enable us to explore our widget prototype variants with participants. Our driving principle was to maximize ecological validity as much as possible while balancing this with generic teleoperation tasks that can generalize across real-world applications. Our testbed includes a real robot and physical space to support navigation and specific actions for participants to complete. We expect that this testbed can also be used by others in the community when they explore teleoperation interface designs.

We target our testbed, and thus our exploration, to the specific case where a person

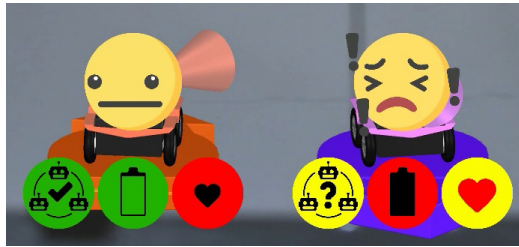


Figure 5-8. Our team state representations for two team robots with emotional information encoding. Note that the message state is embedded onto the face.

teleoperates a single robot from a desktop interface, using an egocentric live video feed from the robot (Figure 5-1). Following, the participant needs to monitor the state of one or two additional robots, which are assumed to be autonomous, using our widget proto-

types (Figure 5-8). We acknowledge that this limits our exploration specifically to single-robot egocentric robot control (in comparison to, e.g., overview command interfaces). However, we note that this remains a primary standard interface for robot teleoperation.

Our testbed has a navigation task where participants must navigate the robot between waypoints, action tasks where participants need to perform operations while displaying other robot team members' states in real-time on the screen (Figure 5-9).

5.2.1. Robot Team-member Simulation

Robot team-members were fully simulated, including their general navigation, gaze direction, and overall states. The robot team-members were moving in the virtual world related to the operated robot – the robot controlled by the user is always located in the virtual world's origin. The simulations updated the team robots' states (movement vector, current position, other robot states) 30 times per second. We generated the simulations using keyframes that consist of timestamps, current robot states, position, and direction information. Instead of randomly generating robot states, we generated keyframes to maintain consistency between participants: all participants saw the same robot states and changes in the same order and timing.

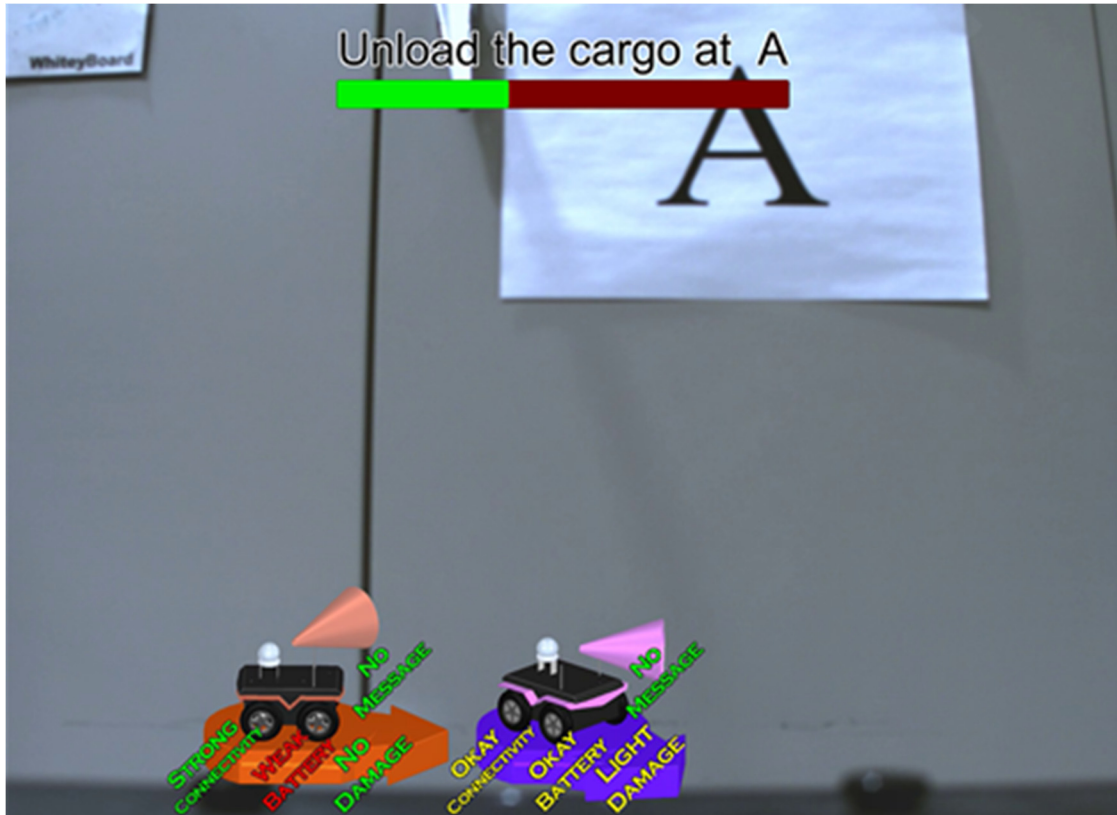


Figure 5-9. A participant is unloading the robot. The mission goal and its progression are shown at the top center of the screen. Note that while the goal is always displayed, the progress bar only appears when the robot is at the designated location.

5.2.2. Navigation Task

We provided participants with a top-down map of the space (Figure 5-10, displayed on a secondary monitor) and told them to navigate to a waypoint shown on screen. We set six designated destinations in the space with a label from A to F. We placed a physical copy of a paper sign for each designated destination on a wall or obstacle in the room. We do not show the robot's position on the map to require the participant to engage in the spatial navigation task, including localization and orientation while avoiding obstacles. If robot collisions make significant changes to its space (e.g., moving a table), the on-site researcher would quickly fix it while the robot was away from the obstacle.

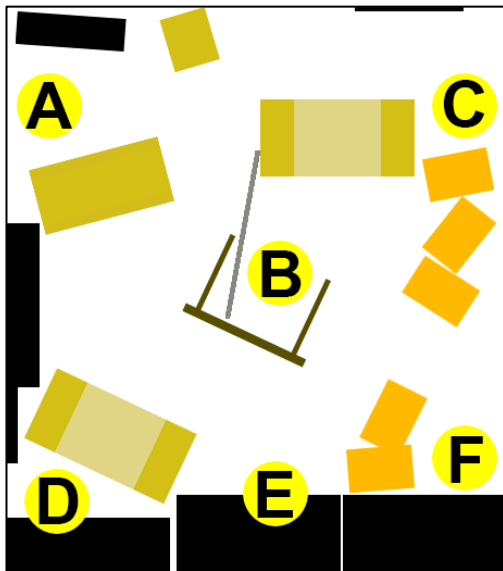


Figure 5-10. A top-down map of the environment. The map shows two underpasses (desks, lighter shading), obstacles, and designated locations. The map is always on the secondary monitor during studies, but the robot's current location is not marked.

5.2.3. Action Task

Our goal for the task was to require engagement from the participant but not be technically challenging in any way specific to our robot or task. As such, we created simulated actions that only required participants to press a button rapidly, like a video game.

Our two simulated tasks were loading and unloading cargo from the robot. Once the robot arrives at a designated way-point, the required task is displayed on

screen (loading or unloading), and a progress bar is displayed (Figure 5-9). When the participant hits a button, the progress bar increases. However, it also decreases with time, requiring the participant to perform work actively to complete the progress bar. Our fill rate is adaptive, filling more per button press as time passes, to avoid the task being too easy or too difficult and to ensure that all participants can complete the task at some point. Upon completion of the task, we asked the participant to navigate to the next designated location. We designed this task to provide some challenge, but not to be frustrating, and provide a sense of achievement once completed.

5.2.4. Measurements

Our testbed's primary purpose was to evaluate how well our widget designs enable people

to maintain awareness of other robot states while teleoperating a mobile robot. Thus, we decided against the common practice of providing paper questionnaires at predictable times. Participants could prepare for them by focusing on the widgets and even refer to the screen to answer questionnaires – this would measure legibility, not ongoing participant awareness.

Our solution was to build a questionnaire interface into our control interface. At

unpredictable times, the interface screen would blank (hiding the screen and widgets), and a questionnaire would appear (Figure 5-11), thus measuring ongoing awareness. We detail the specific questionnaires themselves in Section 5.3.4 and 5.3.5 as they evolved throughout our design exploration.

5.2.5. Implementation

Participants sit at a PC with two monitors, a 28-inch UHD monitor for our teleoperation interface and a 24-inch FHD monitor to show the map. A participant controls Jackal using a gaming joystick (Figure 5-12, right). We positioned the monitor about 65cm away from the participant's eyes, although we did not strictly control this distance, and participants could lean in or away. We placed the keyboard and mouse on a tray under the desk and the joystick on the desk. However, we allowed participants to change the initial desk set up for

[illegible]

Figure 5-11. Our on-screen questionnaire to measure a participant's team state awareness. We did not include the color coding in the questionnaire. When the questionnaire pops up, the teleoperation interface blacks out its screen, stops all processes, and waits for the completed questionnaire.

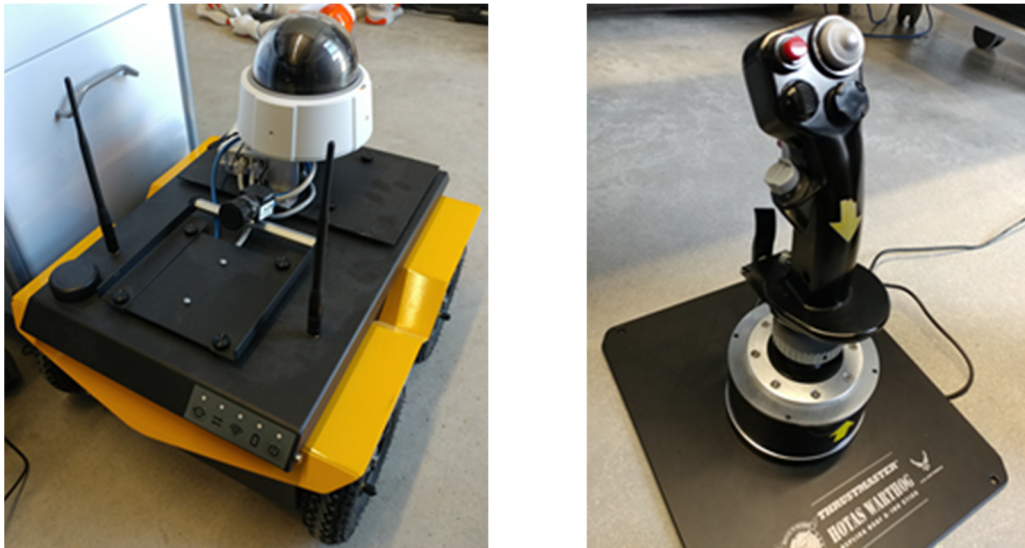


Figure 5-12. Jackal UGV robot from Clearpath Robotics Inc. (left) and ThrustMaster HOTAS Warthog Flight Stick (right) used in our design explorations.

their convenience. Participants use the keyboard and mouse to complete the on-screen questionnaires.

We implemented our interface using Unity3D. Our robot, a Clearpath Robotics Jackal UGV robot (Figure 5-12, left), is 43cm wide, 51cm long, 25cm tall, and about 17 kg, running on ROS indigo. We used a Point Grey Flea3 FL3-U3-13E4C-C with Tamron 1:1.4 8mm ϕ 25.5 lens for the remote video.

5.3. Design Exploration

We conducted a 5-stage design exploration with variants of our widget prototypes and evaluation design. Given our exploratory aims to learn about a range of design parameters broadly and iteratively, we selected a methodology to evaluate several variants with low numbers of participants. We aim to gain high-level feedback and insights across a range of features rather than definitive study results with more participants on fewer features.

The joint-faculty research ethics board at the University of Manitoba approved this study. We recruited participants from our general university population, with each participant receiving \$15 for their time. We conduct our design exploration using the testbed detailed in Section 5.2.

5.3.1. Exploration 1

This initial exploration aimed to investigate our primary overall variables. Firstly, we wanted to compare our new widget designs to the text (base case). Secondly, we wanted to test the animation (vs. no animation) and emoji (vs. no emoji, just icons). Further, this serves as a feasibility test of our testbed and evaluation method. This exploration uses colored assets.

Measuring Operator Awareness of Team-Robot States

Our general approach was to evaluate ongoing participant knowledge of the states of the robots represented by the widgets to test if they supported maintaining awareness. We did this primarily by asking participants to recall the specific robot states. Figure 5-11 shows the initial pop-up format: it asks participants to report on all elements we incorporated into the widget design. Note that this questionnaire only targets the one-team-member robot case.

We also asked participants to report their confidence in their score, to measure how much they felt they were guessing, and to explain the robot state in their own words, to see if they had a general awareness even if their exact state reporting was not correct.

General Questionnaires

Before the test, we administer a demographics questionnaire that collects participant age, biological sex, 3D video game skill and play frequency, vehicle driving skill and how often they drive, and whether they have previously participated in a study with robots.

After completing the teleoperation tasks with a particular widget design, we collected participants' self-reported level of nausea, sense of task performance, how much the widget demanded their attention or was distracting and helped them maintain awareness. We also administered the NASA Task Load Index (TLX, Hart & Staveland, 1988) scale to measure self-reported workload and collected open-form written comments.

After finishing all the conditions, post-study, we further collected participant preference on representations and free-form general thoughts on the widget and task.

Qualitative Investigation

We planned to explore participant feedback on the post-condition and post-test written questionnaires to meet our design exploration goal. We conducted open coding on this written data to investigate what worked and did not work for participants, what items were confusing, and explore potential reasons behind performance differences.

Tasks

Participants completed the testbed task (Section 5.2) five times with different robot widget design: base case (text) and all four combinations of animated versus static and emoji versus just icons: animated emoji, animated icon, static emoji, static icon. In this design exploration, we provided all representations in full color.

We fixed each condition at 6 minutes and 30 seconds long excluding the time taken

to answer the pop-up questionnaires, which appeared at predefined intervals (but unbeknownst to the participant); we did not mention the number and timing of questionnaires to participants. We had sufficient tasks that all participants would require more time than allocated, enabling us to end the experiment after the same duration for all. The pop-up robot state inquiry questionnaire appeared three times per each condition.

Procedure

After the participants arrived, we introduced the idea of teleoperation. We motivated the need for an operator to control one robot while monitoring others, generally explaining the challenge of minimizing the required cognitive load. We administered the demographic questionnaire and provided a training session on how to control the robot using the joystick. We further introduced the concept of monitoring the other robots in the team during training, the states that the robots are communicating, and our widget designs to convey those states. We explained the teleoperation tasks and gave an example of the look of the pop-up questionnaire. We provided ample time for the participant to learn the interface and widgets until they indicated they were ready to move forward.

Participants completed the testbed task with all five conditions, with the order of conditions counterbalanced using an incomplete Latin square. Before each condition, the researcher reminded the participant of the condition and the requirement to maintain awareness of the team robots, and provided an opportunity to ask questions.

After each condition, we administered the post-condition questionnaire. At the end of the study, we administered the post-test questionnaire (see Appendix C for questionnaires).

We recruited 15 participants from the general university public.

Results

Our qualitative analysis revealed the primary theme was that participants tended to comment on how easy a technique was to understand. Five of the 15 people indicated that they felt the text method conveyed meaning precisely, “*words are a much much more clear [sic] interpretations comparing with icons*” – P7 and “*I don’t need to think [with TEXT]*” – P12. This feedback corresponded to the same participants rating the text as the one they liked the most.

In contrast, others (10 out of 15) noted that the icons and emojis require less effort than text: “*the text representation involved mentally deciphering the words and the color code associated with those words*” – P1 and “*I preferred the icons over the text because I found it easier to look at a picture representation than to read words while trying to complete a task*” – P9. This feedback also matched the participants who did not list text as their preferred method.

Many people (7 of the 10 who did not prefer text) noted that the emoji facial expressions added confusion: “*I prefer the simple icons over the ones with character, because for me the ones with character was confusing me*” – P6 and “*the face is not clear and it is very big*” – P11. However, three of the 10 did find that the emoji provided an overall view on the state: “*with character makes it easy to understand the overall situation of team member*” – P4 and “*because it [STATIC EMOJI] is easily to recognize and receive by specific characters, colors, than words description*” – P5.

Many people (7 of the 10) preferred the static representations over the animation. Negative comments toward animation were “*static can help in avoiding distractions*” – P4 and “*the animated icons made the screen too busy so that I was feeling stressed and*

overwhelmed and couldn't focus on the task at hand as well” – P9. There were some positive comments toward the animations: “I prefer the animation over static because animation grabbed my attention better than with just static. Also, you can see which state it is in in the corner of your eyes while trying to finish the given tasks” – P6.

For our quantitative analysis we note that given our small sample size (15), we have limited ability to draw concrete conclusions from our statistical results. We conducted ANOVAs to test the impact of our widget designs on our measures.

First, we conducted one-way ANOVAs with planned contrasts, comparing our new widget designs (animated emoji, animated icon, static icon, and static emoji) against the text (base case). We found no effect of interface on recall of robot state, participant confidence, or nausea. We found an effect of widget type on TLX performance subscale (“How successful were you in accomplishing what you were asked to do?”, $F_{4,56}=2.99$, $p<.05$, Figure 5-13). Planned contrasts revealed that participants felt they performed worse with static emoji ($F_{1,14}=4.78$, $p<.05$, $M=7.87/20$, $SD=5.14$ – lower numbers indicate better result) than text ($M=5.53/20$, $SD=4.47$). We did not find any other effects.

We further conducted two-way ANOVAs, excluding the text case, to investigate our two design dimensions; animated vs. static by emoji vs. icon-only. As above, we found no effect of interface on recall of robot state, participant confidence, or nausea. We found a primary effect of emoji (emoji vs. non-emoji) on participant reported TLX performance ($F_{1,14}=5.92$, $p<.05$): people reported that they performed worse with emoji ($M=7.07/20$, $SD=4.06$ – lower numbers indicate better result) than non-emoji ($M=5.00/20$, $SD=4.04$). We did not find any other effects nor interaction effects.

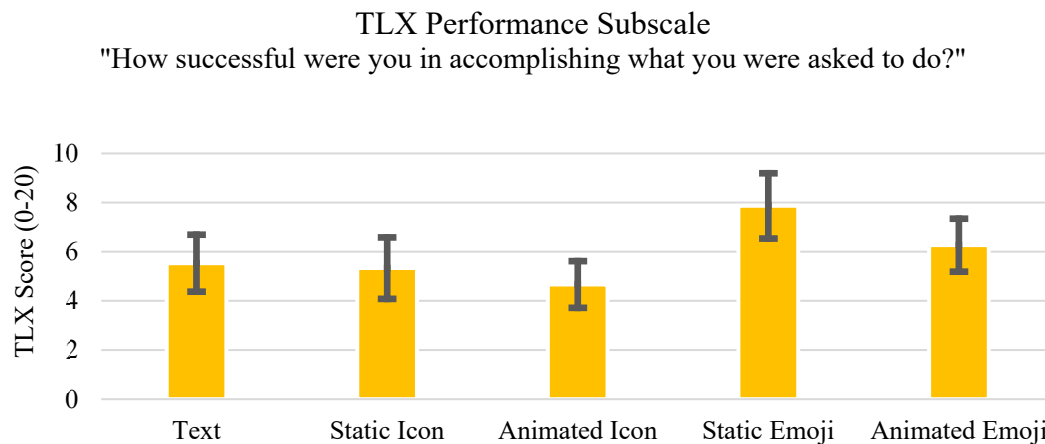


Figure 5-13. Exploration 1 NASA TLX performance subscale report. Error bar indicates the standard error. Lower scores are more positive. That is, the participant thought that they were more successful.

Lessons and Next Steps

Our qualitative data generally supports our design approach, that most participants seem to find the icons less mentally demanding than the text. However, about a third of the participants felt otherwise, and thought that text was a reasonable approach. Our inconclusive quantitative results do not provide further insight, as no interface seemed to perform better than any other.

Regarding animation, participant consensus did seem to be that the distraction of motion was a detriment, although some noted the benefits of pulling attention. Further, despite some participants noting the emoji benefits, both the qualitative and quantitative data suggest that people may find them confusing and more challenging to interpret overall.

Our concern with the data was the apparent lack of performance results. How well participants could recall the robot states – we have no evidence that any design performed better (or worse) than others. Looking at the data, we first noted that the robot was continually moving and looking around, making it somewhat difficult for participants to answer

those state points. However, reconducting our analysis without these measures did not yield results. Further, we found that while the overall average accuracy was 63%, some participants managed to obtain 100% recall, suggesting a potential ceiling effect on our measurement. Of particular note is that we have no evidence of text yielding worse performance than the other methods, despite our expectations.

Figure 5-14. Our on-screen questionnaire to measure a participant's team state awareness for two team robots. See Appendix C for the high-resolution captures.

To follow, we decided to test the same experimental design with more robots to increase the difficulty and avoid the ceiling effect while simultaneously investigating the impact of monitoring more than one robot team member.

We also note that some participants noted the effect of color, which was one of our design approaches; we re-visit this in a later stage.

5.3.2. Exploration 2, Two Team Robots

We conducted a second exploration to investigate the impact of requiring participants to monitor two robots, with two related on-screen widgets (Figure 5-8 and Figure 5-9) and the expanded questionnaire (Figure 5-14). We followed the same procedure otherwise as above, Exploration 1. We recruited five pilot participants and compared new results to the first exploration.

Results

Our qualitative analysis uncovered a clear difference relating to the task difficulty, where some participants complained of the complexity, which was not raised in the same fashion in the previous study: *“there are too many things to look out for at the same time”* – P21. Given the low participant count, we did not conduct statistical testing. However, the average recall performance dropped by 10% (from 63% to 53%), and the overall NASA TLX scores increased from average 7.87/20 to 10.93/20. Despite the lowered state recall accuracy (moving away from the potential ceiling effect), we still did not see clear performance difference between widget designs.

The text approach also did not perform more poorly, as we expected it might. One participant linked this to our performance measurement instrument: *“it is easy to remember while answering the questionnaire”* – P21.

Other results echoed what we found in Exploration 1, where participants praised the use of color coding, *“same color represents same status is very good”* – P23 and others generally commenting on the icons, *“icon is easy to understand and simple and static can make me easy to find the moving and looking around information”* – P23. One person also noted how the animated icons drew positive attention: *“it was easier to interpret without focusing too much”* – P24.

Lessons and Next Steps

Our results indicated that the second robot did impact participant perceived task load and performance on the state recall task (with the caveat that we do not have statistical signifi-

cance). However, despite this we did not see a difference in performance across our interface designs, with text still not performing more poorly, contrary to our expectations. However, one participant provided a potential hint, noting that the text in the questionnaire matches that specific interface design.

In considering our participant feedback, we noted that color may provide a memory cue and that participants may monitor the color (and changes) using their peripheral vision. This observation further motivates the exploration of color versus greyscale variants.

5.3.3. Exploration 3, Greyscale variants

We conducted a greyscale pilot with five new participants to compare against our color results to date. Our greyscale variants force participants to interpret the icon or text data instead of relying on color only. We maintained the same procedure from Exploration 1, which only included one robot to monitor.

Results

The overall results of the greyscale study matched what we found in the previous explorations. The overall recall accuracy was 59%, with TLX load reported as 10.78/20 on average.

Participants prominently compared techniques against text, noting the ease of interpreting the icons, *“because it is less demanding on the visual processing, has little to no distraction, gives room for better mental representations for the individual, which is absent in textual forms”* – P30. With two particularly noting the benefits of text: *“the words allowed me to understand the icons more quickly”* – P28. Finally, feedback on the emoji was generally negative: *“emoji was hiding the movement of the robot”* – P29.

Lessons and Next Steps

We found similar results with the greyscale version as with the color; the minor differences found (in the direction of greyscale being slightly harder) does not suggest that we need further inquiry involving more participants and statistics. Notably, this result does not support the idea that participants may be relying heavily on the color coding only for monitoring robot states. We still believe that the color coding is beneficial for our motivational reasons, supported by qualitative feedback. These results suggest that we do not need to investigate further.

Based on participant feedback focusing on text, particularly participants making an explicit link between our widgets (with text shown) and our evaluation instrument (where they select the same text), we noticed a potential confounder in our study design. For example, P21 preferred text representations over others because it was easy to remember while answering the questionnaire. In the text interface case, participants match the words from the widget to the questionnaire. In contrast, in the icon cases, they need to remember the icons' meanings and translate the text on the questionnaire. This matching may provide an advantage to the text case and explain our results with text performing similar to other widgets, as we expected text to perform more poorly. To investigate this, we update our evaluation instrument for the next exploration.

5.3.4. Exploration 4, Improved State Questionnaire

Drawing from our realization of the text-based questionnaire being a potential confound, we re-designed the questionnaire to match the widget interfaces' designs. That is, we kept the text-based questionnaire for the text widget case (as already shown, Figure 5-11). We




The robot is:

☐ Moving ☐ Not Moving




☐ Looking Around ☐ Not Looking Around

The robot has:




Connectivity

☐  ☐  ☐ 

Battery

☐  ☐  ☐ 

Damage

☐  ☐  ☐ 

Message




☐  ☐  ☐ 

Figure 5-15.A




The robot is:

☐ Moving ☐ Not Moving




☐ Looking Around ☐ Not Looking Around

The robot has:




Connectivity

☐  ☐  ☐ 

Battery

☐  ☐  ☐ 

Damage

☐  ☐  ☐ 

Message




☐  ☐  ☐ 

Figure 5-15.B

Figure 5-15. Our modified questionnaire with the icons that match the widget designs. We show a portion of the questionnaire for icons only (A) and emojis (B) for convenience. See Appendix C for the high-resolution captures.

modified the questionnaire for the icon and emoji conditions to show the specific items to match the widget designs (Figure 5-15). We removed any state-specific information (e.g., color, facial expressions for emoji) to ensure they are generic.

We followed the same procedure as in previous explorations, emphasizing participant training for answering the questionnaires to ensure they understood the new additions. We kept the full-color widgets but used the two-widget case (to monitor two robot team members) to avoid potential ceiling effects as noted earlier.

We recruited 10 participants to test the matching option questionnaire; the additional amount (compared to 5 participants for previous phases) enables us to do statistical exploration.

Results

We start with our quantitative results to focus on the impact of the new evaluation instrument approach. We first used one-way within-subject ANOVAs with planned contrasts to compare our new widget designs against the text case. We did not find any statistically significant interface effect on any measure taken, including recall performance or self-reported workload. The average performance was 42% for correctly answering team robot states, and the average TLX response was 12.05/20, which is higher than previous explorations. People performed worse comparably and thought that they were too loaded to accomplish the given tasks.

We conducted two-way repeated-measure ANOVAs to investigate the impacts of our animation and emoji dimensions (animated vs. static, and icon only vs. emoji), excluding the text case. We found a main effect of animation on the TLX effort dimension (“How hard did you have to work to accomplish your level of performance?”). Participants reported that the animated interfaces ($F_{1,9}=7.57, p<.05, M=15.05/20, SD=3.85$, Figure 5-16) required more effort than the static ones ($M=13.75/20, SD=4.22$). However, according to Figure 5-16, it appears that the weak performance of the emoji animated may be driving the result.

Our qualitative feedback again echoed prior studies. Comparison with text was an overall theme, with half of our participants preferring the text method, e.g., because “*ICON is hard to memorize. I honestly prefer the color not the texture itself*” – P34 and “*it is much more clearer in sending the message*” – P35. Those who did prefer the icons noted that it is “*because it’s easy for me to see and does not distract me from the task*” – P31.

Again, as previously, participants noted the difficulty with the number of states,

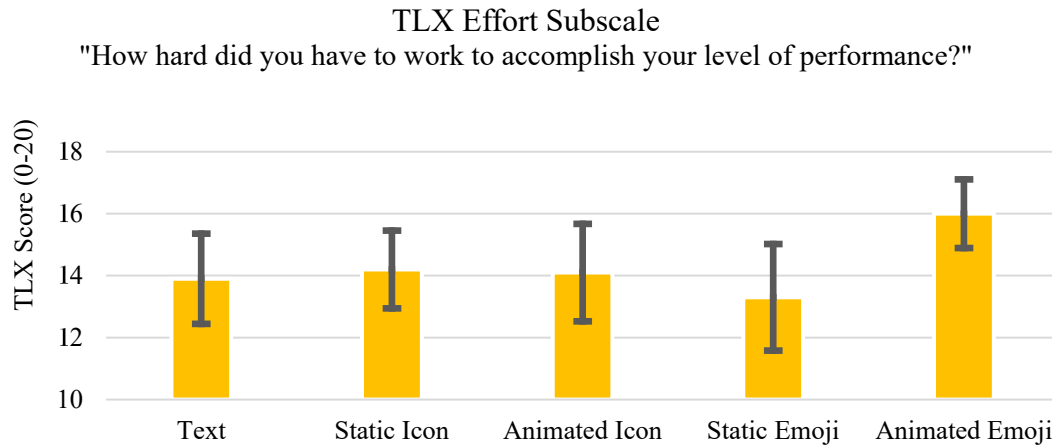


Figure 5-16. Exploration 4 NASA TLX effort subscale report. Error bar indicates the standard error. Lower scores are more positive.

e.g., “two team members are too much. I could remember only one of them” – P40.

Lessons and Next Steps

We did not see an improved performance or change in results despite our improved questionnaire design; however, we found a stark decrease (although we did not conduct statistics on this change). We note that earlier indications of high workload were echoed here, with participants noting the difficulty of keeping track of everything. As such, for the last exploration phase, we again attempt to resolve the problem of evaluating the operator’s awareness in our scenario. We dropped this questionnaire and explored a slightly different approach to measuring awareness of the robots’ states.

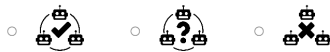
5.3.5. Exploration 5, Alternative Questionnaire

To get a more accurate measure of how well participants can maintain awareness of robot states using our widgets, we made significant changes to our protocol.

First, instead of having our pop-up questionnaire ask about the entire robot state at

Regarding your left team member. . .

How is your left teammate's connectivity?



In general, what are your left teammate's states?



Figure 5-17. Our stripped-down questionnaire, which only asks one state of one team robot and the robot's general well-being.

once, which adds a time delay for the person to recall by the end of the questionnaire, we reduced the pop-up to only inquiring about one aspect of the robot state at a time (Figure 5-17). It would only ask about one of the robots at once,

and inquire about either connectivity, battery level, damage, or message. In addition, we asked for an overall sense of the robot's well-being. We removed questions about moving and looking around; as previously mentioned, the robots are always moving or looking around, so these were less valid.

Second, we aimed to collect more data to compensate for the reduced detail of the pop-up questionnaire and increase our result's precision. Given the strong indication in earlier studies of animation being distracting, we only kept the static conditions for this iteration. We did keep the emoji despite negative results, as given the novelty of the approach, we wanted to collect more data. Thus, we have three conditions: text, static icon, and static emoji.

As a result, we extended the task time from 6 minutes 30 seconds to 10 minutes and increased the number of pop-up questionnaires from 3 to 8, to collect more data. We recruited six people to participate in this study.

Results

The overall accuracy dropped to 21% (from 63% in Exploration 1 and 42% in Exploration 4), with TLX score averaging 10.15/20. We gained no additional insight from qualitative analysis.

Discussion

It appeared as if our new state measurement technique reduced participant accuracy dramatically (from 63% in Exploration 1 down to 21%), with TLX scores remaining about the same as the other explorations. Reflecting, we conducted a post-hoc meta-analysis of participant performance scores across our evaluation methods. That is, we compared our original pop-up questionnaire that always showed text states, against our revised version that showed icons, and against our final reduced questionnaire; this meta-analysis only included the text, static icon, and static emoji cases to enable us to include Exploration 5. We note that this is a post-hoc analysis across studies, which involved other variables. As such, we do not make statistical claims but rather use this to inform our future inquiry. A one-way ANOVA reported an effect on state report accuracy ($F_{2,38}=29.84$, $p<.05$), where participants reported more accurately with the initial, all-text questionnaire ($M=61.5\%$, $SE=2.47$) than the matching questionnaire with icons ($M=41.7\%$, $SE=3.90$) and the final more straightforward questionnaire ($M=20.8\%$, $SE=5.04$).

5.4. Guidelines for Graphical Widgets Representing Robot

Team Member States

Here we summarize our overall findings resulting from our five-stage design exploration. Overall, all representations we created helped people maintain team robot states' awareness to a similar degree, with no obvious winners or losers. The unexpected primary finding was how well the textual representation fared.

Because our sample size is small in each exploration, however, we cannot make a

strong claim with our quantitative results. As we understand this limitation, we are relying primarily on qualitative feedback from participants (written comments) to discuss design insights for team state widgets. Below, we synthesized five design insights for team teleoperation interface designers.

text is a viable candidate – short one- or two-word text state representations performed as well as icon representations. While one may assume that text is slow to read, perhaps with a short text it can become similar to an icon and easily recognizable (iconification). For example, P12 mentioned that they do not need to think with Text (5.3.1.Results) and P28 mentioned that they could understand text much quickly than icons (5.3.3.Results).

people feel icons are easier – more than half of the participants reported that they prefer icons and emojis, despite a lack of apparent performance increase. For example, P1, P9, P23, and many others explicitly mentioned about it (introduced across multiple subchapters). According to their comments, they felt icons were easier to understand. Consider using icons in cases where people's perception of workload is essential.

anthropomorphic representations may not be clear – while some enjoyed the faces, most participants reported that the emotional encoding information was not clear, and in some cases distracting. For instance, P6 said that they preferred the simple icons over emojis because the emojis are confusing (5.3.1.Results).

animation: balance distraction with attention grabbing – participants generally reported that animation could be distracting (e.g., P4, 5.3.1.Results), although a clear minority found them attention-grabbing in a positive way (e.g., P24, 5.3.2.Results). This supports prior work (Rea et al., 2017) on balancing distraction.

color is useful to show the level of robotic states – participants found color coding to help

maintain awareness of team robot states, as the colors distinguish the states' level (severity or urgency). For example, P34 explicitly mentioned that they preferred color coding while disliking both text and icon representations (5.3.4.Results). We recommend teleoperation interface designers to use color coding for the level of robot states if applicable.

5.5. Discussion and Future Work

We designed novel visual widgets to represent generic robot states in a specific kind of teamwork teleoperation scenario. Overall, the widgets performed acceptably well, and through our exploration, we identified the core challenge with exploring and assessing how such widgets could help operators maintain awareness.

In addition to all our designs performing reasonably well, the unexpected result was that the text case performed better than our anticipation. One potential explanation can be iconification, where the short text becomes an icon of sorts: our text state has color and two simple words.

Another interesting overall finding was that participant thoughts on their performance, as noted in qualitative results, does not correlate to their actual performance or their TLX report. It will be necessary to explore further what impact this participant perception may have on overall task performance. Also, to measure load we could extend our existing method, for example, by using reaction time (Dubé & McEwen, 2015; Sternberg, 1969) or pupil dilation (Jiang, Zheng, Bednarik, & Atkins, 2015; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004).

We note that the opinions on our animated icons were mixed. One problem is that, in our animated cases, the icons were always animated. An alternative approach would be

to have the icons animate only when the state changes, such that it draws attention only when necessary.

Our emoji icons did not perform well overall. Our initial design approach was to create an emotional encoding that may leverage the person's social interaction system, supporting recall. However, perhaps the level of abstraction introduced by our emoji created confusion. This limitation and further investigation on leveraging social interaction without creating vagueness is essential for future work.

Overall, we felt that our testbed was a success, despite the challenges we faced with evaluating awareness. This measurement problem is not solved, and additional methods need to be investigated. Further, the testbed could be improved by incorporating the team robot states (e.g., whether a robot is damaged) into the participant's required teleoperation actions. Currently, they are unrelated. Making the other robot states more critical for the primary task may add motivation to maintain awareness of those states.

5.6. Summary

We presented a set of novel widget designs and prototypes to convey the general states of the robot team members in a teleoperation scenario. Further, we developed a novel and generally applicable team teleoperation scenario testbed that others can use for similar work.

We presented the results from a design exploration where we investigated the impact of visual design parameters and explored the challenge of evaluating awareness of each robot's state. This investigation resulted in a set of guidelines for graphically representing robot team member's states that can aid future work in the area.

Our work is directly related to team teleoperation, where an operator dedicates one robot in the team using an egocentric teleoperation interface. While our team state representations can apply to multi-robot teleoperation scenarios where a single operator manages multiple robots, in such case, the operator needs a different set of information. In the next chapter, we explore how interface designs can support an operator's awareness of robots when switching their focus between robots.

Chapter 6 Control Transition: Utilization of Visual Transition as an Information Source in Multi-robot Teleoperation

When an operator manages multiple robots in multi-robot teleoperation, the operator encounters situations where a robot needs the operator's direct attention, for example, due to unexpected situations (Rosenfeld et al., 2016) or their lack of permission for making an important decision (Wright et al., 2018). The operator must perform a *control switch*: switching their focus from their current task (either controlling a robot or working on non-teleoperation tasks) to the robot, which requires the operator's direct attention. We call this process a *control transition*, where a control switch happens (Figure 6-1), and the operator successfully transitions their focus to the new robot.

As the operator worked on other task(s) before the control transition, they may not have enough situation awareness to complete tasks using the new robot after the control transition. Therefore, the operator must reconstruct situation awareness before accomplishing their mission goal successfully and efficiently with the new robot.

When an operator switches to a new robot, they need to understand a great deal of information. For example, the new robot's current states, what the environment (where the new robot is) looks like, how the environment changed over time while the operator was not paying attention to it, what the robot was working on, and how the robot was performing their tasks. We pose a question: how we can support the operator to get the information and understand them. This problem of supporting an operator's situation awareness is an ongoing challenge and open research problem in the field of human-robot interaction (Phillips & Jentsch, 2017; Roldán et al., 2017; Seo, Rea, et al., 2017).

We see an opportunity to transition to the new robot (i.e., the control transition) when an operator transitions their direct attention and control from one robot to another to communicate with the operator about the new robot's state. We want to use the control

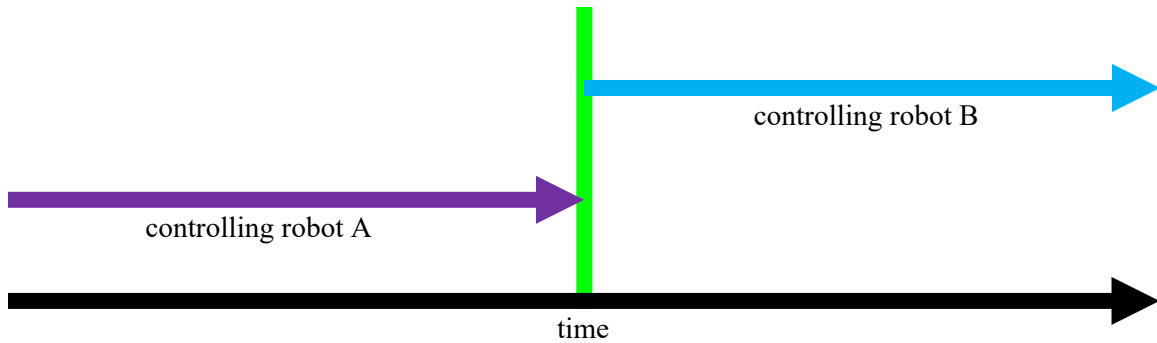


Figure 6-1. The concept of the control transition in multi-robot teleoperation. At the current point on the time domain (green vertical har), the operator switches their control from one robot to another. That is, robot A is no longer under a control of the operator after this point.

transition moment to provide the robot’s relevant information to the operator instead of merely changing the interface information. We call this *informative visual transition* as we provide information to the operator during the user interface’s visual transition (teleoperation interface) to help them reconstruct their situation awareness.

How to design this transition, however, is not trivial. Our methodology was to look to other fields, where managing transitions is the key to success, and survey their techniques to explore the feasibility of the informative visual transition in multi-robot teleoperation. We recognized a similar problem in film and various forms of media. A film must carefully orient audiences to a new scene and situation to ensure they can follow the story arc. For example, movie may have a scene where the camera pans or uses the dissolving effects of two scenes to provide a transitional effect (Katz, 2019). Other media also has their own transitions. Novels use narrative exposition as a scene transition (Hill, 2016). Video games use a loading screen to simulate scene changes (Figure 6-2, left) and to provide the next scene’s context information (Figure 6-2, right).

Our fields of exploration are human-computer interaction (transitions and notifications of software changes), cinematography (transition and camera techniques), and video games (information presented during data loading). We collected various techniques that

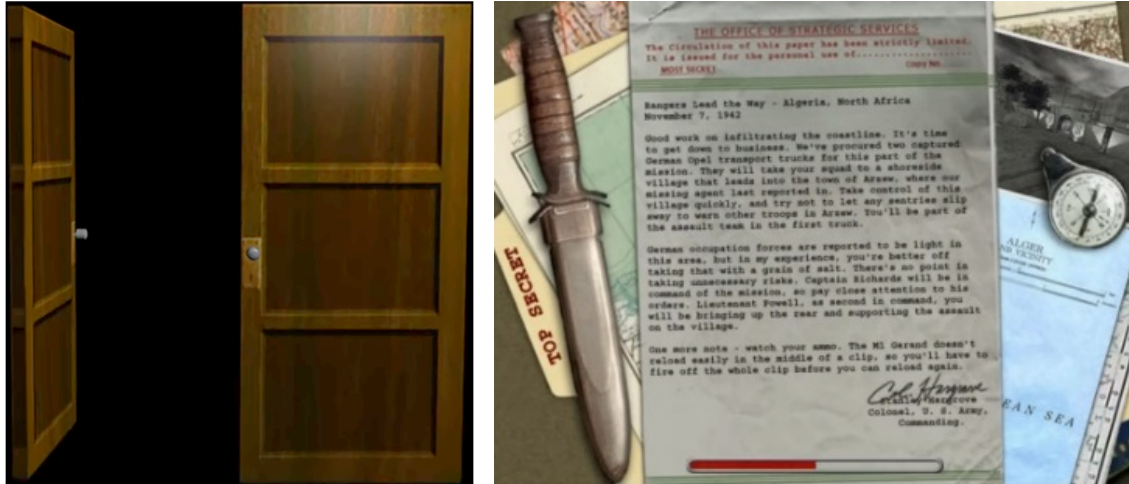


Figure 6-2. Example loading screens from games. Left loading screen provides a tip of moving from one space to another using the door animation (*Resident Evil*, Capcom, 1996). The right loading screen gives a gamer the next stage's context information (*Medal of Honor: Allied Assault*, Electronic Arts, 2002).

provide information to their audience/users and discussed how they could provide information for multi-robot teleoperation. This discussion serves as an initial design framework of informative visual transition for multi-robot teleoperation interfaces.

In the rest of this chapter, we survey techniques from other fields, present our analyses of these techniques, and culminate in a framework for informative visual transition in multi-robot teleoperation. We present a problem explanation and initial sketch of a survey in this chapter along with preliminary results. Our contribution in this chapter is the discussions of the problem and the approach. That is, we bring vocabularies from other fields' techniques to human-robot interaction and collections of ideas to improve multi-robot teleoperation interfaces. We have focused on providing a novel design paradigm (informative visual transition) to interface designers to improve future teleoperation interfaces.

6.1. Problem: Information Required during Control Transition

During multi-robot teleoperation, operators need a lot of information after a control switch

to reconstruct their situation awareness and control the new robot. Specifically, they need to understand what the robot is for and what task it should perform. That includes the robot's location information (e.g., in the nuclear reactor or near a pile of debris), physical configuration (e.g., having an end effector of gripper or screwdriver), and its embodiment (e.g., flying robot vs. ground robot).

Reconstructing situation awareness can also be found in other domains, for example, farming business and medical fields, which share some similarities. We took example situations, envisioned what information could be useful for their situation awareness in these domains, and explained how they are related to multi-robot teleoperation. While it is different from scientific analyses, this exercise can help us realize how having information within sight could increase people's situation awareness. We would also like to discuss what kinds of information can help situation awareness after task switching to initiate our thought process for supporting the operator's control transition.

Afterward, we took our initial list of information and discussed with human-robot interaction experts to identify what types of teleoperation information can help the operator understand the situation after the control transition. From the discussion, we extracted the idea that the operator may require three significant items, which are transition notice, event history, and current states.

6.1.1. Task Switching Examples in Other Domain

Switching people's focus from one task to another happens daily within real-world application domains, including farming, shopkeeping, office work, and so on. To quickly ac-

comply with people's goals and determine their next moves after switching their tasks, similarly to multi-robot teleoperation, they need to know what happened before (i.e., history) and what the current situation looks like (i.e., up-to-date state information).

For example, some North American farmers use driverless tractors to increase their productivity. When a farmer uses an autonomous tractor, they can work on their other business-related tasks (e.g., accounting). However, if the tractor stops due to an unexpected situation, the farmer must quickly resolve the issue to maintain their productivity. If the farmer can learn the cause of the stop (e.g., animal in the path, mechanical failure, or software error), the path that the tractor moved on until now (to decide new path), and the maintenance history of the tractor (to identify mechanical problems of the tractor) then they may be able to resolve the situation quickly.

There are more examples. A doctor in an emergency room quickly diagnoses a patient and moves on to another patient. As they jump into a case, a nurse collects the patient's medical history (e.g., currently taking medicines, diseases the patient suffered in the past, and their family's medical history) and briefs the current status to the doctor. Additional information helps the doctor treat the patient correctly (e.g., avoiding allergic reactions to medications). During the transition between patients, the nurse's information can help the doctor decide treatment(s) for the new patient.

Task switching in restaurants occurs very often for servers. Especially, when there are many customers in restaurants, servers are busy switching tasks between taking orders, bringing orders to customers, working as a cashier, cleaning tables, and handling unexpected situations (e.g., customer spills waters all over the place). A server's task performance relies on how accurately and quickly they can recall the previous tasks' states (e.g.,

which orders have and have not been delivered to tables). In the long term, the success of the restaurant is tied to the server's task performance.

In all these cases, people can benefit if they have the required information on hand after task switching. That is, people can quickly reconstruct their situation awareness, get to their new task, accomplish their goals if they have the information of what they must do next (i.e., knowing transition), what happened before (i.e., history), and what the current situation is (i.e., current states). In the next subsection, we check if we can attach teleoperation situations to the categories and discuss what information helps the operator reconstruct situation awareness.

6.1.2. Information for Reconstructing Situation Awareness in Multi-robot Teleoperation

Similar to the real-world examples, an operator in multi-robot teleoperation can quickly reconstruct their situation awareness after a control switch from one robot to another if they understand the necessary information. That is, the operator must know that they are controlling the new robot to reduce their mode error. The operator must understand the situation of the new robot in the past where they were not paying attention to (i.e., the working history of the new robot). In addition to the history, knowing what the operator must do and can do with the new robot is important to successfully perform their tasks. This summarizes that the necessary information for the operator during the control transition is (a) what they must do next (i.e., knowing transition and control the new robot), (b) what happened before (i.e., history around/of the new robot), and (c) what the current situation is (i.e., current states). We visualized this concept in Figure 6-3.

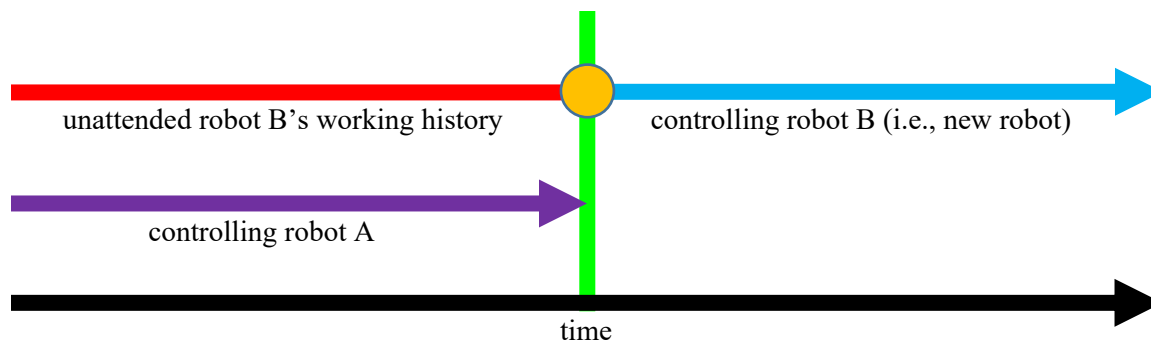


Figure 6-3. The concept of the control transition. Green vertical bar conceptualizes the current point in time when the operator must switch their control from robot A to robot B.

- Robot A was controlled by the operator, but it will not be after this point.
- Robot B was doing its work without the operator's attention in the past, but it will be controlled by the operator.
- For the operator to perform their tasks with robot B successfully, they must know that they are controlling robot B from now on (transition feedback), comprehend the robot B's working history (red bar), and understand current states (orange ellipse).

Picking the three types of information, we held discussion sessions with human-robot interaction experts to compile the details of each information category to help the operator achieve their situation awareness after the control transition. While we do not claim that the list is a complete set, we believe that this can help us anchor our further exploration on informative visual transition in multi-robot teleoperation.

Knowing Transition – Transition Notice (Feedback)

An operator needs to know that the control transition is happening to prepare themselves to reconstruct their situation awareness with the new robot. The teleoperation interface can provide visual feedback to the operator to notify the control transition.

In user interfaces, the feedback has an important role: it helps the users comprehend the system's status (Saffer, 2013). For the control transition during multi-robot teleoperation, when the operator initiates the control transition, the feedback helps them know that the interface took their command. If the system initiates the control transition, the feedback

helps the operator be ready for the transition. Thus, in either case, the visual feedback (i.e., transition notice) can help the operator know that their control is switching and reduce any accidental mode errors.

Event History

As discussed in Chapter 2, current situation elements' perception is an essential part of having situation awareness (Endsley, 1995). The same applies to reconstructing situation awareness. After the control transition, the operator needs to have the perception of elements (i.e., the state information of the new robot and the robot's environment) to reconstruct their situation awareness.

In addition to the perception of *current* elements, which is in the next subsection, the robot's working history can help the operator understand what the robot was doing. Additionally, the environment history (e.g., wall collapsed near the robot) may explain how the robot ended up calling for the operator's help. For example, suppose the robot was navigating to one place but failed due to unexpected environmental changes. In that case, the history of environmental changes helps the operator understand why the robot cannot reach their destination and plan how to overcome the situation. These information are inclusive – thus, we call this *event history* all together.

There is different history information that helps the operator have situation awareness around the new robot. As an autonomous robot is working on its own, the robot logs its actions, sensor readings, and other information (i.e., the event log), which can be essential to diagnose the robot's problems and challenges. This information includes the nature of the task, how and when the robot attempted it, when the robot completed it, and any other notable events that complete the event history.

The event history helps the operator understand the past progress toward the robot's current task, and in some instances, to set the next plans accordingly. For example, if the operator knows the robot's path in search and rescue, the operator can focus on the area where the robot has not been through to find any anomalies.

Current States

In addition to the event history, the current information is essential for the operator to reconstruct their situation awareness. Various up-to-date states of the robot and the environment help the operator determine what they can do next. For example, the robot's configuration explains what the robot can do (driving/flying, grasping an object, etc.). The gyroscope reading tells the robot's physical orientation (flipped or not). The information regarding the surrounding environment provides a hint of the robot's assigned tasks (office environment or nuclear reactor).

Robots are customizable (i.e., configurable) by swapping parts. For example, the same robot can have different end-effectors such as nippers or suction cups depending on their goals and tasks (e.g., cutting wires or climbing up on a glass wall). With heterogeneous robots, which has different configuration or embodiment, an operator may get confused what the new robot can do after the control transition. Thus, by providing each robot's details during the control transition, the operator can quickly decide what they can do.

Mobile robots maintain a large set of internal states, including connectivity (e.g., connections to the base station or between robots to share sensor readings), battery level (if not tethered), inertial readings, gyroscope readings, servo positions (e.g., its kinematics, i.e., current servo configuration/position), and so on. The operator can quickly achieve some level of situation awareness and perform a task efficiently by knowing some of these

states at the right timing. The robot's kinematics can be useful when an operator must go through a tight passage. Since the information varies from robot to robot, we must provide at least some (if not all) of this information to the operator during the control transition.

An operator might control a robot to resolve situations that robots cannot solve by themselves. However, without knowing the robot's challenge and the situation, the operator may have to spend a long time to figure out the problem. Thus, a summary of the robot's current challenges can help the operator understand the situation and speed up the operator's decision, leading to overall success.

The robot's surrounding information, the robot's current geographic location and the remote environment can help an operator understand the current scene. The operator can plan their next move according to environmental information. Then, the question is how we can provide the information we iterated and discussed above.

6.2. Survey of Visual Transition Techniques in Other Fields

Regarding teleoperation information (*transition notice, event history, and current states*), we initiate a discussion on presenting the information to an operator during a control transition, such as maximizing the use of informative visual transition. We found similar problems and solutions in other fields. The media and literature fields make specific efforts to carefully orient audiences to a new scene and situation to ensure that they can follow the story arc. We first gather transition techniques from other fields, including film and video games. We develop a framework that aims to provide example techniques from other fields (e.g., film, video games). Our framework provides vocabularies and keywords that explain how and why the techniques are relevant and useful to teleoperation.

6.2.1. Camera Techniques in Cinematography

Film scene transitions happen by blending the visual effects of the two scenes. There are primarily seven effects and many varieties derived from them: cut, fade-in, dissolve, white-in, wipe, white-out, fade-out (Katz, 2019). Visual effects, however, are ambiguous in terms of their meaning. For example, fading out after an actor's death conveys a different feeling than fading out while two people laugh together. Because of this reason, we cannot explicitly classify each effect to a specific meaning. Rather, we need to admit that the meaning will vary depending on the highly artistic expressions in film.

Even so, one common element regardless of their meaning is that these visual effects provide the transitional effect. That is, whether the fading effect used to convey sad feeling or happy feeling, they show the visual transition in the end. For this reason, we marked visual effects as transition notice and extended our survey to camera techniques.

Along with visual effects, one way to orient the audience's atmosphere in film is camera technique. There are many known camera techniques and their improvements in films. With novel hardware and knowledge (as well as a director's artistic talents), these camera techniques keep evolving, and cinematographers introduce new techniques every day. For this reason, it may not be practical to list all existing techniques.

Further, we could not find academic publications regarding camera techniques in cinematography, possibly because they aim toward practical applications. Therefore, we picked a list from an educational web article (the title is *Film Studies 101* ... Freer & Gibbs, 2018) as our initial survey chose and grouped the techniques for quick introduction below.

Instead of simply listing the techniques, we tried to understand how to apply them in film, how they were used in popular movies, and what the directors communicate using

them. We iterated our classification multiple times; we do not claim that this is the complete or perfect classification. Some techniques can be used for different purpose like the visual effects are. However, this exercise and classification help us anchor our discussion in designing informative visual transition and possible uses in multi-robot teleoperation.

Transitional

The following techniques are often used in story transitioning either when a part of a story is completed (e.g., an episode is finishing by tilting the camera upward) or when a portion of a story is completed (e.g., a person finishes talking and zoom focus is changed to the background of the person).

- High Angle Shot: the camera moves away to the high angle, i.e., toward the sky while taking the character on the ground (e.g., *The Shawshank Redemption*, 1994, at the end of escaping the prison)
- Tilt: the camera changes its angle upward or downward (e.g., *Robert Altman's Nashville*, 1975, at the end of the movie, the camera points upward to the sky)
- Zoom: the camera stays where it is (i.e., the camera lens does not move) but shifts the focus to another subject (e.g., *The Conversation*, 1974, opening park scene)

Intensify

The following techniques provide an idea that something is not right: dreamlike, unsettle, surreal, and so on.

- Dolly Zoom: the camera takes a subject or an environment while shifting the camera's physical position as well as zoom level at the same time to providing the distortion of the reality (e.g., *Jaws*, 1975, when Brody realizes abnormality at the

beach)

- Dutch Tilt: the camera is physically tilted to convey an unsettled feeling (e.g., *Mission: Impossible*, 1996)

Sequential

The following techniques are used to show events and happenings in the order of their occurrence. These techniques can keep the audience's attention while quickly moving the story forward.

- Bridging Shot: stitching multiple scenes translucently (e.g., *Indiana Jones and the Raiders of the Lost Ark*, 1981, the scene with an airplane, a map with red lines, and a set of event scenes)
- Whip Pan: quickly swiping scenes after scenes with wind-blowing-whip audio effect (e.g., *Hot Fuzz*, 2007, the primary character's achievements in the opening of the movie)
- The Sequence Shot: the camera keeps following a character (e.g., *Touch of Evil*, 1958, the 3 min 20 secs opening in one-go)

Exploration

The following techniques are used to show the environment where the story happens.

- Aerial Shot: the camera flies around the primary character (e.g., *The Sound of Music*, 1965)
- Establishing Shot: the camera flies over buildings with narration introducing the place (e.g., *The Shawshank Redemption*, 1994, at the beginning of the movie)

- Top Shot: like establishing shot, but the camera takes the ground from the perpendicular angle in the sky (e.g., *Taxi Driver*, 1976)
- Pan: the camera pans in 360-degree to capture the environment around it (e.g., *Brian de Palma's Blow Out*, 1981)

Immersion

While every camera technique makes the audience immerse into the film, these techniques provide an immersive feeling toward a character or an environment.

- Handheld Shot: the camera is held by one of the characters in the scene (e.g., *Scorsese's Mean Streets*, 1973, the fighting scene in the pool hall)
- Point-of-view Shot: taking scenes from a character's perspective (e.g., *Doom*, 2005, after the primary character is awoken, and *Hardcore Henry*, 2015, the entire movie)
- Locked-down Shot: the camera is locked at its position like a surveillance camera (e.g., *Manhattan*, 1979)
- Tracking Shot: following a subject be it from behind or alongside or in front of the subject (e.g., *Paths of Glory*, 1957)

Emotional Attachment

Films use many techniques to make people emotionally attached to fictional characters. The following techniques are examples of using a camera to enhance emotional attachment for the audience to the characters.

- Arc Shot: the camera circles around the subject while keeping its focus on the subject (e.g., *De Palma's Carrie*, 1976, the couple dance)
- Close up: taking the full face in the frame (e.g., *The Passion of Joan Of Arc*, 1928)

- Deep Focus: having a sharp focus on the foreground, middle ground, and background (e.g., *Citizen Kane*, 1941, where a child is playing outside in the background, while the child's mother discusses the future of her child with other gentlemen)
- Over-the-shoulder Shot: taking a character over other character's shoulder (e.g., *The Godfather*, 1972)

Observing Characters

These techniques are primarily for the audience to observe characters and environments in a film.

- Cowboy Shot: stereotypical cowboy duel scene (e.g., *The Good, The Bad and The Ugly*, 1966, gun duel scene)
- Long Shot: taking an entire character or object from head to foot (e.g., *Lawrence of Arabia*, 1962)
- Low Angle Shot: taking characters from the low angle to show authoritative figures (e.g., *Matrix*, 1999, the scene with two agents)
- Medium Shot: taking a scene with the subject's upper body, facial expression, and actions (e.g., *The Searchers*, 1956)
- Two Shot: taking two people in the frame (e.g., *Magnolia*, 1999)

Supportive

These techniques are used in a supportive manner.

- Crane Shot: shot using a crane (e.g., *Gone with The Wind*, 1939)
- Library Shot: shot pulled from a library (a.k.a., stock shot, e.g., *Tarzan*, 1958)
- Matte Shot: foreground action with a background, painted onto glass or computer-

generated graphics (e.g., *Planet of The Apes*, 1968)

- Money Shot: money flies from the sky (e.g., *Independence Day*, 1996)
- Steadicam Shot: using a hydraulically balanced camera (e.g., *Goodfellas*, 1990)

Summary

We note that there are numerous camera techniques; our list of camera techniques can be serving as a steppingstone for coming and future discussions. We would also like to note that some techniques can be combined with others and used for other purposes. Again, our goal is not to have the perfect list. Instead, we are looking for inspiration from the media field to utilize control transition as a visual information source in multi-robot teleoperation.

6.2.2. Video Games

Like cinematography and their camera techniques, video games have also evolved and innovated over several decades and continue to grow. Many scene transitions in this field help players follow the story arc, add visual aesthetics, or pass the time waiting for data to load. Like the camera techniques, there are numerous scene transitions and techniques in video games. In some cases, this can be even complicated because of their interactive nature. Thus, we limited our exploration to the inevitable transitions, loading screens.

Hardware is still bottlenecking the video game world because loading takes time. While hardware advancement reduces data loading time significantly for video games from the same era, people desire more advanced graphics and an improved gaming experience with more advanced hardware. These lead to more data being required in the next generation of video games. This race is on-going: hardware advances, peoples' desire increase,

video game data enlarges, and the cycle repeats.

To move from one scene to another in video games, due to the volume of video game data (e.g., graphics texture, audio, etc.), loading the data from storage to working memory and unloading the unnecessary data from working memory are inevitable tasks. Game developers are developing highly efficient programs, and game designers are working toward making more entertaining loading screens.

In this work, we specifically look at game designers' efforts on loading screens and adapt the techniques to our media-inspired informative visual transition framework. Unfortunately, we could not find academic references regarding loading screens in video games. Thus, we referenced journal articles and opinion videos.³ Then, we put the video games that we know and classify their loading screens based on their characteristics.

Things to Read

Games provide stories, lore, and gameplay tips during their data loading.

- *Medal of Honor: Allied Assault*, Electronic Arts, 2002. (Figure 6-2, right).
- *Dragon Age: Inquisition*, Electronic Arts, 2014.
- *Middle-earth: Shadow of Mordor*, Warner Bros. Interactive Entertainment, 2014.
(audio flashback of the story with gameplay tips)
- *Just Cause 3*, Square Enix, 2015.
- *Assassin's Creed Odyssey*, Ubisoft, 2018.

³ URL: (youtube.com/watch?v=RSV4rHCPJ0M), (youtube.com/watch?v=hhVT7ydgGxo), (youtube.com/watch?v=hhVT7ydgGxo), and (gamesradar.com/the-secret-art-of-the-video-game-loading-screen-and-why-they-wont-be-going-away-anytime-soon/)

Session Statistics

The games provide summary statistics of a previous session(s).

- *DOOM, id Software, 1993.*
- *Rocket League, Psyonix, 2015.*

Hints for Location Change

The games provide a hint of location or scene change.

- *Resident Evil, Capcom, 1996.* (door opening)
- *Mass Effect, Electronic Arts 2008.* (elevator scene)
- *Dragon Age: Origins, Electronic Arts, 2009.* (moving on the map, Figure 6-4)
- *Destiny 2, Activision, 2017.* (the planet at the end of the loading)

Interactive

Players can play a mini-game or a training session to practice the game mechanic.

- *Assassin's Creed, Ubisoft, 2007.*
- *Bayonetta, Sega, 2009.*
- *FIFA 19, EA Sports, 2018.*

Simple Progression

An object or a progress bar is indicating the loading. Some games provide artwork or audio feedback.

- *Hexen: Beyond Heretic, Id Software, 1995.*
- *Battlefield 1942: The Road to Rome, EA Games, 2003.*
- *XCOM 2, 2K Games, 2016.*

Summary

Many other games are not listed but fall into one of these categories. We also note that the list is not perfect. The primary question, related to our work, is if these techniques from video games are applicable for the control switch during multi-robot teleoperation (not creating an exhaustive list of loading screens and their classification) and how we can apply these techniques to control transition during multi-robot teleoperation. Hence, we continue our exploration instead of perfecting this list.



Figure 6-4. A part of a map in a video game with the player's trail. The trail starts at the highlighting circle on the top, passes the sword icon (a battle scene), and continue stretching downward. Screenshot from *Dragon Age: Origins*, BioWare, 2009.

6.2.3. Techniques in Other Fields

We touch on various techniques from other fields in this section to remind readers of many more techniques. While we did not include an exhaustive list of these in our framework, as discussed later, we note that the knowledge of more such techniques could be helpful in future exploration.

Expositions in Literature (Transitional Phrases)

Transitions in literature are essential anchors in the story. With good transitions, readers can follow the story with minimal effort; however, writing good transitions is not trivial. Additionally, “[t]ransitions can be as short as one line, or they can be paragraphs or

pages long” (Hill, 2016). This definition implies that an author can decide to have a short phrase or another full story as an exposition in literature. We do not use literature exposition examples in this work, simply because they are not visual and not applicable in teleoperation interfaces. However, we briefly discuss the impact of simple transitional phrases from writing resources as a reminder of existing transition techniques.

Transitional phrases are often grouped based on their logical relationships between the former and latter sentences, including addition contrast, and emphasis (Odegaard Writing & Research Center, n.d.; Student Writing Support, n.d.). Using these relationships, a reader can expect how the story will continue. A simple transition is a tool that briefly notifying a reader of the transition. For example, when they read “in addition,” they expect additional information on the theme of the previous sentence, or with “but,” they expect a contrasting story to the previous one.

While literal transitional phrases and expositions are not directly applicable to visual transitions in teleoperation, this may be of interest or inspiration for other researchers and visual designers. However, our goal is to develop a framework inspired primarily by film camera techniques and video game loading screens. We develop the framework before continuing our exploration in literature transition techniques.

Visual Transitions in Graphical User Interfaces

Visual transitions happen in all graphical interfaces. When the content is updated (e.g., media players), new information is displayed (e.g., notifications), existing information is collapsed (e.g., menu items), or an application is brought forward of others (i.e., switching the application focus). We can find many examples throughout current applications, from operating systems to task-specific applications.

A critical role for visual transitions in graphical user interfaces is to minimize the loss of context for users. For example, when users change their focus from one window to another, the operating system animates the changes while showing the other ones (e.g., Microsoft Windows Aero). Modern websites provide other examples, where clicking a tab activates animated scrolling instead of an instant page update.⁴ In addition to keeping the context, these animations act as feedback to the user's command.

Providing information during a visual transition is also found in many research projects while not being explicitly mentioned. For example, researchers used techniques to provide visual information (Kamaruddin & Sulaiman, 2018), including animations (e.g., shrink one camera feed while expanding the other in switching camera feeds Singh et al., 2013) and transitional supports in multi-robot teleoperation (e.g., Calhoun et al., 2012; raper et al., 2008).

Despite its existence and a set of examples in both industry and research, the idea of utilizing visual transition as an information source is not fully developed in multi-robot teleoperation. Therefore, we develop an initial framework that introduces a novel design paradigm for the control transition in multi-robot teleoperation.

6.3. Our Design Framework: Connecting Control Transition

Problem and Visual Transition Techniques

We develop an initial design framework that provides the operator's required information

⁴ smashingmagazine.com/2013/10/smart-transitions-in-user-experience-design/

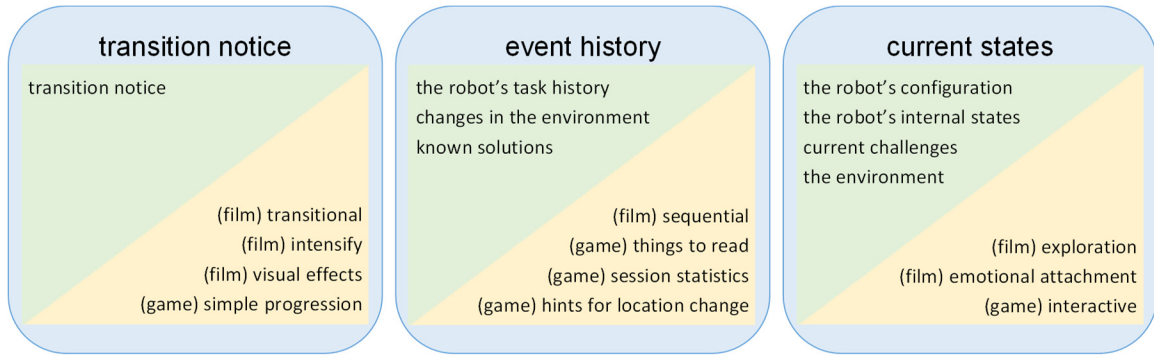


Figure 6-5. Our initial design framework. The blue box contains the type of information, the green contains teleoperation information, and the yellow contains groups of techniques taken from other fields.

to reconstruct their situation awareness after the control transition and the visual techniques we collected from other fields. This design framework focuses on anchoring our discussion on a small scale and discussing our novel idea's possible implementation, informative visual transition in multi-robot teleoperation. This interim framework has a lot of room to be evolved and expanded, which we leave as our future work. Our current framework shows what techniques can convey particular type of information (Figure 6-5).

6.3.1. Categorization of Media Techniques

Among the collected techniques, we selected techniques used to show progression (simple progression game loading screens), visual effects (effects in cinematography), and emphasis on a scene (transitional and intensify camera techniques). Designers can use these techniques to show changes off the screen, acting as a visual transition notice. We listed the techniques included in this category:

- (game) simple progression
- (film) visual effects
- (film) transitional

- (film) high angle shot
 - (film) tilt
 - (film) zoom
- (film) intensify
 - (film) dolly zoom
 - (film) Dutch tilt

We note that other techniques can also be used to show visual updates in user interfaces. However, we selected simple techniques that we think can be primarily used for transition notice in multi-robot teleoperation interfaces in this category.

For event history, the interface must display rich information to the operator. In other words, the contents may contain a lot of detailed text, statistics, or visual information. In that sense, the below techniques may be suitable to show the rich information:

- (game) things to read
- (game) session statistics
- (game) hints for location change
- (film) sequential
 - (film) bridging shot
 - (film) whip pan
 - (film) the sequence shot (showing the robot's work session)

To visualize the current states during the control transition, we choose techniques used to explore characters and environments. For example, the game's interactive loading screen allows players to explore possible movement and explore other features. This can turn into an exploration of the remote robot's current states and its physical configuration.

The list techniques of our choice are:

- (game) interactive (exploring current states by interacting with a virtual robot)
- (film) exploration
 - (film) aerial shot
 - (film) establishing shot
 - (film) top shot
 - (film) pan
- (film) emotional attachment (as to show the robot's current exterior condition)
 - (film) arc shot
 - (film) close up

Our selection and categorization can serve as an anchor point for design discussion for informative visual transition in multi-robot teleoperation. Despite our effort, we admit that this is a proof-of-concept and requires further improvement. Hence, we call the initial framework, which requires many iterations ahead. We leave the improvement as future work and focus on the potential of our framework.

6.3.2. Initial Design Discussion Using the Framework

We emphasize that this work focuses on exploring a novel interface design paradigm in multi-robot teleoperation, instead of creating a perfect framework. Our main goal is to bring up the discussions of informative visual transition in multi-robot teleoperation and how we can utilize other fields' techniques to design teleoperation interfaces. In this note,

we take existing multi-robot teleoperation interface cases and initiate design discussion using our interim framework. We check if our proof-of-concept framework can help explain the information load of visual transitions and brainstorm control transition designs in multi-robot teleoperation through this exercise.

For this quick evaluation, we surveyed existing multi-robot teleoperation interfaces. We collected the information available to us. However, due to its proprietary nature, there is limited access to the interfaces. Many teleoperation interfaces are in-house custom-made for specific robots and tasks. They are either not released to the public (e.g., interfaces used in research projects) or only available with purchasing robots (e.g., robot control packages in the robotics industry). That is, our explanation regarding the interaction of the interfaces is limited. Further, multi-robot teleoperation interfaces with a full egocentric perspective are rare in the market despite their benefits. Many interfaces often use a manager view with additional information widgets.

We grab available screenshots of existing multi-robot teleoperation interfaces. Regardless of their layouts, we presume that the interface visualizes the control transition and discusses informative visual transition designs we could make using our design framework. We note that our purpose is not to underrate these interfaces (they have been designed and built by professionals over several years) – we use them as a tool to discuss control transition designs and check if our framework can help drive the discussion about multi-robot teleoperation.

Case 1: Multi-UAV Testbed Interface

The paper (Calhoun et al., 2012; Draper et al., 2008) introduced a multi-UAV interface used to control semi-autonomous UAV robots (Figure 6-6). The interface presents various

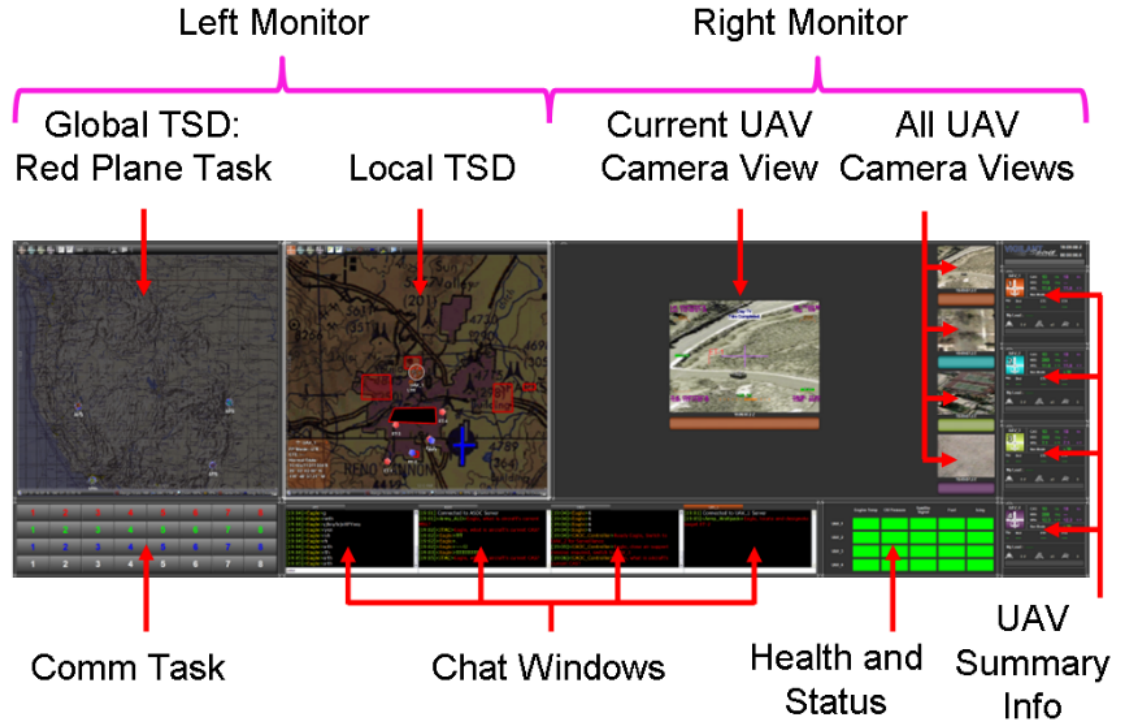


Figure 6-6. Multi-UAV testbed interface (Calhoun et al., 2012; Draper et al., 2008).

information about current states to the operator. While the interface is mostly presenting current states, some event history (logs) stays there for the operator to read again. However, we do not think the interface is mainly providing history in depth.

The interface consists of global and local maps, robot information, and the currently selected robot's camera view. Especially in (Calhoun et al., 2012), the authors experimented on three different transition conditions for switching camera views of two geographically separated UAVs. The three conditions are (1) discrete switch (the camera view switched immediately to new camera view), semi-autonomous (fly-out the camera, provide a global visual view and wait for the user input to fly-in), and autonomous transition (fly-out the camera, provide a global visual view, wait for two seconds, and fly-in to the next robot's camera view automatically).

While the authors discussed the operator's performance and situation awareness

using different transition techniques in the paper, they did not focus on the information gain through different visual transitions. For example, in our initial framework introduced above, we mentioned that visual effects could be used as feedback (change of focus) to users, or aerial-shot camera technique can present current robot states.

We attached visual transitions to what information we can provide to the operator, while the paper did not discuss this. In our opinion, for the three transitions introduced in the paper, in their discrete switch, there is no additional information provided (transition is quicker than the other two, but no additional gain for situation awareness). In contrast, the user can gain additional information in the semi-autonomous and autonomous conditions (may take time to finish the transition) and achieve some level of situation awareness.

Our framework can explain the semi-autonomous and autonomous transitions: the virtual global view serves as Establishing Shot or Top Shot in the camera techniques that show the environment from a top-down view and help users have the geographical knowledge. By observing the environment, the operator can gain information on the current states, including a brief overview of the environment, the robot's location in the global environment, and the geographical relationship of robots compared to their peers. The information collectively supports the operator's situation awareness, leading to improved task performance.

Case 2: RoboLeader User Interface

Our next case is the RoboLeader interface presented in J. Y. C. Chen & Barnes, 2012a, 2012b; J. Y. C. Chen, Barnes, & Qu, 2010 (Figure 6-7). Like Case 1: Multi-UAV Testbed Interface, the interface primarily presents current states with some of each robot's geographical location history on the map. We do not think that there is any hint for control



Figure 6-7. RoboLeader multi-robot teleoperation interface (J. Y. C. Chen & Barnes, 2012a, 2012b; J. Y. C. Chen et al., 2010).

transition other than highlighting the thumbnail of camera views.

The interface consists of a primary camera view (selected robot's camera view), thumbnails of all robots' camera view, mission intelligence and objectives, and a global map. We do not have information regarding how the interface reacts to the control transition between robots. However, we believe that taking one of our framework techniques can help the operator understand various information.

For transition notice, regardless the initiator of switch control between robots (i.e., whether the operator initiates control switch or the system), any techniques we listed in the transition notice category can help the operator notice the changes. In particular, if robots have similar scenery (like the robot camera view thumbnails in Figure 6-7), the operator needs some noticeable visual feedback to reduce the mode error. However, given the use

of the interface in the military, we think that quick and straightforward transitions are preferred, such as visual effects in cinematography.

Our framework can continue helping the discussion regarding providing additional information during control transition. For example, for any event records (intelligence shown at top-left corner in Figure 6-7, the interface can visually remind the operator of any updates near the newly monitoring robots using a Bridging Shot technique or an interface inspired by the video game's session statistics loading screens. Further, when the operator switches their control between robots, they can be reminded of the robot's current states and goals, for example, using a virtual arc shot to provide a look-around the robot.

Case 3: SMP Robotics Multi-robot Teleoperation Interface

The last interface we surveyed is from SMP Robotics Systems Corp. (Figure 6-8). Unlike others (running on a PC), this particular interface is designed to be running on a tablet. Due to its limited display space, we think they put three different modes (selectable by clicking a button on the screen's top-center). We do not fully know their available features, but we assume their appearance on the official website's screenshot.

In the interface, a list of robots and the event logs on the left, the information panel for the selected robot and the command list on the right, and the map and visual feeds in the middle. From the interface's overall view, we assume that control transitions between robots are initiated by the operator (by touching the robot on the map or the list). Once the robot is selected, the information on the right panel would be updated. From the control transition (i.e., selecting a different robot), the operator would expect the changes, since they are the one who initiates the transition.

We initiate design discussions using our framework. When the operator switches

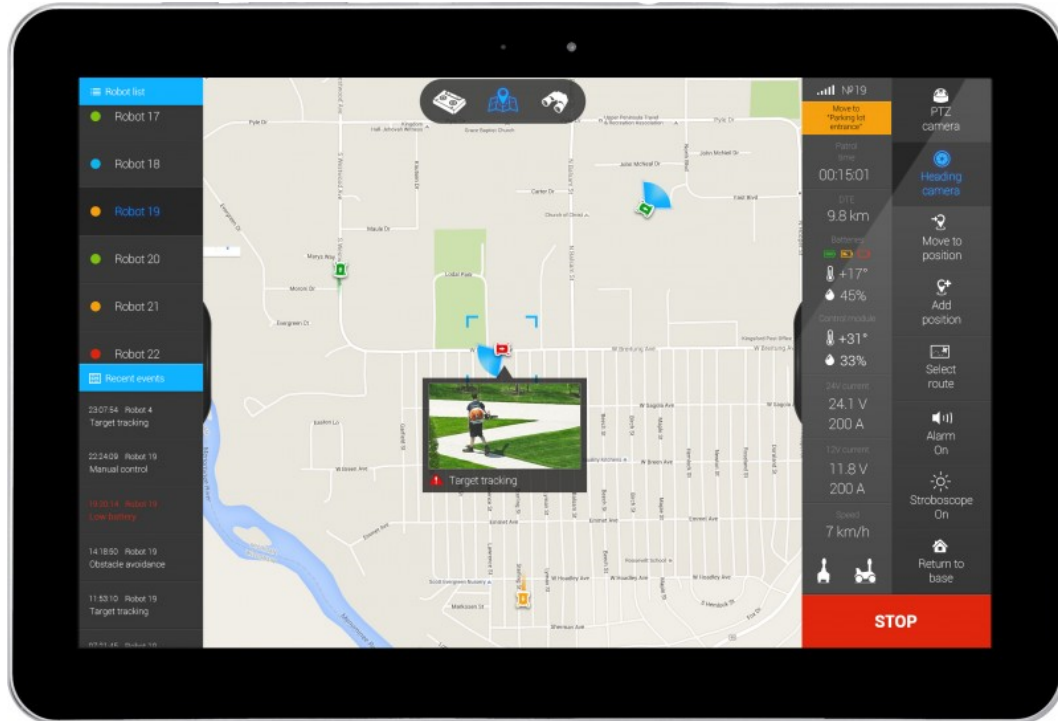


Figure 6-8. A multi-robot teleoperation interface by SMP Robotics Systems Corp. Retrieved image from smprobotics.com/products_autonomous_ugv/multi-robot-control-system/ on April 20, 2020.

their control between robots, visual cues can help the operator have awareness of the remote environment and the newly controlled robot. For example, when the control transition happens, visual cues can help the operator recognize the transition or complete it. Another example is that film's visual effects such as dissolve can be directly applied to the map. As the dissolve of the map's viewports (the current position and the newly controlled robot's location) happens, the operator can be ready to read new data from the map. Camera techniques (transitional – tilt) can follow the interpolation between the current location and the robot's location to highlight the updated location information and the environment. Once all the visual updates are complete, the operator understands that the transition happens. We believe that during the transition using our example designs, the operator can attain additional information (e.g., location changes).

As the interface already shows multiple robots' event records on the left panel (e.g.,

target tracking and low battery), the operator can achieve situation awareness around the robot quickly by highlighting the newly controlled robot's history. From our design framework, a designer can use the techniques listed under the *event history* category. For example, an additional overlaying widget with a statistical summary (like a game's loading screen, e.g., battery level, number of completed waypoints, etc.) can provide a quick summary to the operator. Or a quick replacement of the map to a video stream of significant events (e.g., using film's bridging shot or whip pan camera techniques) can be impactful for the operator's understanding of the robot's task history.

As the control transition completes, the operator must understand the robot's current states on top of how they changed over time. The interface already shows the up-to-date information as well as the available commands on the right panel. However, without visual cues, the operator may have to put extra effort into finding the differences. Here are some design ideas for visual cues using our framework.

We can replace the map with a virtual robot view for a short period. With the virtual robot, the interface can use the arc shot camera technique to show the robot's exterior with visual marks on its physical damages. While doing so, it can also highlight each part of the virtual robot to indicate available commands (e.g., a game's interactive loading screen which provides tips of new moves or commands). Or instead of the virtual robot, the real robot can pan its camera to show the remote environment's video feed, which can help the operator understand the remote situation quickly.

Our design ideas are directly taken from our design framework and updated the information to be conveyed to the user.

6.4. Summary

This chapter discussed a novel interface design paradigm to provide additional information during the control transition in multi-robot teleoperation. We developed a design framework inspired by media and literature techniques that can help control transition design discussion. We introduced possible use of our work and potentials in discussing multi-robot teleoperation interface designs through the exploration exercise.

Despite our effort, the results we presented in this chapter is preliminary. To fully utilize our informative visual transition idea, we need to make a further discussion on this topic, including collecting required information automatically and systematically, choosing the appropriate visualization techniques, and balancing the transition duration and information visualization.

Our framework is yet to be improved – there are many more techniques we can learn in other fields. There can be detailed oriented design guidelines, and we can conduct design workshops to test our framework’s validity. We received approval from the joint-faculty research ethics board at the University of Manitoba (Appendix D); however, we decided not to conduct the planned workshop due to the worldwide pandemic. We leave them as future work while noting that it is essential to develop novel design paradigms for future teleoperation interfaces. That is, we are initiating the discussion and suggesting the novel paradigm for providing teleoperation information to the operator during the control transition in multi-robot teleoperation interfaces. We hope that this triggers further discussion and investigation on novel ideas in informative visual transition.

Chapter 7 Epilogue: Our Work, Significance, and Future Work

This Ph.D. work explores teleoperation interface designs to help an operator improve situation awareness and increase their overall task performance. In previous chapters, we have introduced a selected set of teleoperation challenges: visualizing multiple camera streams in one feed, providing team-robot location information using visual in-feed indicators, exploring various visualization components for multi-robot internal states, and discussing designs of informative visual transitions between robots. While these projects do not cover the entire set of teleoperation challenges, we explored the breadth-based interaction problems in teleoperation scenarios, implemented novel interface designs, and investigated their use through user studies. Our main approach is to collect the remote information and present them to the operator in a way that they can comprehend in time. We found that our approach is useful through formal user studies. In this chapter, we summarize our work, discuss lessons we have learnt, propose future work, and compile our contributions.

This thesis focuses on designing interfaces for operators to improve their situation awareness in hectic remote teleoperation scenarios. We suggested novel teleoperation interface designs that provide multiple camera streams in detail-in-context view (Chapter 3), embed in-feed team location indicators (Chapter 4), present widgets for team robot working states (Chapter 5), and utilize control transition moment as a teleoperation information source for achieving a high level of situation awareness when it is necessary (Chapter 6).

Chapter 3 presents a novel interactive detail-in-context teleoperation interface that integrates a pan-and-tilt narrow-angle first-person view into a wide-angle behind-robot view; operators can move the narrow view around a scene to obtain more resolution when and where needed. Our interface integrates two teleoperation cameras into a single mixed-resolution view, removing the need for an operator to map between the views mentally.

Through a formal study, we collected results that demonstrate our interface's feasibility and show that it can help operators complete search and rescue tasks more effectively than simple solutions.

Chapter 4 demonstrates how in-feed embedded location indicators can be applied when teleoperating a robot with our exploration in indicating team member locations. Compared to a mini-map, a widely used solution for location information, we showed that our in-feed indicators, located on the screen's peripheries, use much less screen real estate, demand less attention, and require less mental mapping. Our results also provide insight into some of the trade-offs between our in-feed embedded location indicators, such as being better for distance or direction information, or which demands more or less operator attention. These results provide a roadmap for future work, continuing to improve location indicator designs using peripherals.

Chapter 5 presents a set of widgets in exploring design strategies for visualizing team robot working states. We aimed to help an operator understand complex robotic team robot states and actions easily and quickly. Our widgets consisted of different information representations such as text, icon, and emotion and graphical visualization parameters such as color, animation, and the number of team robots. We developed a mock team teleoperation scenario and task and on-screen questionnaires to test people's awareness of team robot working states. In the project, we described our reasons for choosing representations and parameters and discussed the advantages and disadvantages of our team robot state representations, which can help interface designers make informed decisions on designing team robot state representations in teleoperation.

In Chapter 6, while we did not conduct a formal study for this project, we suggested

a novel interface design paradigm for utilizing the visual transition moment in multi-robot teleoperation scenarios. Our idea is to fill in the operator's necessary teleoperation information during their control transition, switching their focus between robots. Our design suggestions are inspired by media and literary techniques such as bridging two different views into one shot or providing information during time-consuming loading screens. Despite our framework being preliminary, we discussed how the idea is applicable in teleoperation interfaces, leading to improvements in future designs.

In each chapter, along with the introduction of these projects and their results, we discussed project-specific points such as the monitor's physical size, possible 3D integration of our indicators, and our widget's iconification. Ultimately, we linked our discussions to the operators' understanding of teleoperation information (i.e., the operator's situation awareness). In the following subsections, we link our work to the operator's situation awareness, suggest future work on egocentric teleoperation interfaces, and summarize our contributions to the community.

7.1. Balance Between the Remote Information and the Operator's Workload

Throughout this thesis, we presented more remote information issues: operating robots with increasing numbers of camera feeds and operating robots in groups. In both cases, the operator in search and rescue must have a high level of situation awareness to accomplish their mission. There may be a victim's life at risk. For this reason, many mobile robots

carry various sensors such as cameras, microphones, inertial measurement unit (IMU) sensors, and so on. These sensors are helping the operator expand their perception of elements in the remote environment via teleoperation interfaces.

One dilemma for designing teleoperation interfaces is to determine the right volume of the remote data. While providing more information with increasing numbers of remote sensors may help the operator have the necessary information at hand, it also increases the operator's workload to understand the information. Moreover, without presenting the information in compact (but understandable) ways, it adds time for the operator to check the available data. For example, if two video feeds are presented side-by-side (Chapter 3), the operator must change their focus between views to understand the two views. In other words, teleoperation interfaces should provide as much remote information as possible while presenting them in understandable ways to support and increase the operator's level of situation awareness.

We showed possible avenues in improving egocentric teleoperation interfaces by balancing the volume of the teleoperation information and the understandability of them throughout our projects. Our monocle interface (Chapter 3), for example, shows how we (interface designers) can help the operator maintain situation awareness while reducing their cognitive workload by integrating two video feeds into one feed and presenting their physical relationship. Our in-feed location indicators (Chapter 4) and team-state widgets (Chapter 5) also present all available information in a compact form by abstracting details without losing the core information. We also discussed designing informative visual transitions for multi-robot teleoperation interfaces to increase the operator's comprehension of the remote information when they need the information.

Our work serves as a steppingstone for interface designers to think of abstract data and integrate sensor inputs for more information with simple interfaces. Related areas for future work include the level of abstraction (e.g., balancing between the detail and the abstraction of robot states), visualization of multi-dimension location information (e.g., 3D location indicators for team robots), and integrating strategy for multiple sensors (e.g., merging three, four, or more camera feeds).

7.2. Envisioning the Teleoperation Consequences

For a high level of situation awareness, an operator should project the robot's future status and the remote environment presented by teleoperation interfaces (Endsley, 1995). A simple example is understanding that as the robot is current moving forward, the object beside the robot will be at its back. While this sounds simple, keeping track of all previous states and the updated states can be challenging, especially in hectic teleoperation scenarios (e.g., search and rescue). Many variables are also changing the future status, including the operator's actions (e.g., commanding the robot to move backward), consequences of the actions (e.g., the robot hitting an object), and nature (e.g., earthquake). Thus, the difficulty of continuously maintaining a high level of situation awareness increases in remote teleoperation.

One way to support the operator is to design teleoperation interfaces to keep presenting the state. Our Monocle interface helps the operator keep track of the head-mount camera's gaze direction by showing the outline of the view when it is off to know where it will be when it is on.

Another way is to remind the user of the previous states. Our discussion in informative visual transition pointed to help the operator comprehend the history of the robot's

tasks and changes in the environment, which is ultimately linked to the projection of the near future states.

While we did not explicitly implement or suggest prediction visualization in our design discussions, there are many methods to help the operator project future status via teleoperation interfaces. Teleoperation interfaces might help the operator with prediction of an action's outcomes by, for example, identifying where the robot is about to go (e.g., using augmented reality to draw a line on the floor) or providing an estimation of specific actions (e.g., a ghost image of destination camera position while the camera is moving). However, presenting additional information to the operator may result in negative task performance due to their increasing cognitive load. We leave the detail discussions on the problem, explorations, and implementations as our future work.

7.3. Moving Forward from Our Work

Our work introduces possible avenues for improving egocentric teleoperation interfaces. While the future of egocentric teleoperation interfaces may differ from our expectations, we believe that our ideas advance interface designs. Following, we discuss how we possibly move forward from our current stage.

Some examples of different interfaces include manager views, exocentric teleoperation interfaces, or even interfaces without visual outputs (e.g., speech interaction). Even current multi-drone interfaces are different from traditional egocentric teleoperation interfaces. However, egocentric teleoperation interfaces have their benefits, primarily when the operator cannot rely on the external sensors or cannot physically be with robots. While this interface may get less popular in near future, its importance remains and continued research

on egocentric teleoperation interface designs will have significant benefits.

To retrieve remote information, sensors on robots are essential. As developers keep improving sensors, they tend to become more compact and provide more highly defined data with higher frequencies. Therefore, we can mount more numbers of advanced sensors on a mobile robot of the same size. During remote teleoperation, the increase in remote data retrieved from sensors can help the operator achieve higher situation awareness. However, as we already discussed in Chapter 3, more sensor data (e.g., video feeds) will lead to needing different interface design strategies to present the information in quickly understandable ways.

One such problem that we must address for future improvements of egocentric teleoperation interfaces is their scalability. For example, in Chapter 3, we showed our successful teleoperation interface design that provides two camera streams in the detail-in-context view. When there are two (or more) cameras and when pointing in different directions, the view may not be maintainable in a detail-in-context and instead need camera views separate from each other.

Our widgets introduced in Chapter 4 and Chapter 5 are obviously at their disadvantage regarding their scalability. There will be too many indicators for one screen and be too cluttered if there are many numbers of team robots. Our novel interface design paradigm (providing visual information during the control transition) also suffers from scalability challenges. For example, with increasing numbers of both homogenous and heterogeneous robot team members, the operator must differentiate each robot as they control each one to manage the team of robots. The visual cue alone may increase the operator's mode error (i.e., which robot is under control now).

While we cannot currently provide a direct solution for the scalability, one method to mitigate the problem could be to use advanced artificial intelligent robots. While people will not let robots make all decisions soon (e.g., moral decisions), artificial intelligence can reduce the operator's workload, as they can work without the operator's direct attention. However, we need to explore different interface design strategies with different control schemes to utilize advanced artificial intelligent robots fully.

In addition to the improvements, we can apply to teleoperation interfaces, we need a better toolkit to measure the operator's situation awareness level. This toolkit is essential to objectively determine if the interface helps the operator achieve a high level of situation awareness. In our explorations, we used self-reported questionnaires to measure people's awareness in user studies and collected their task performance statistics (e.g., number of objects found throughout the teleoperation course). While these methods served well enough to compare interface designs, they do not precisely measure the operator's situation awareness. A standard method for measuring situation awareness in teleoperation scenarios will help teleoperation interface designers in the near future.

7.4. Our Findings and Contributions

This thesis attempted to improve interface designs with four different teleoperation scenarios, introduced our solutions (Figure 7-1), and discussed future work derived from our work. Our approach in improving teleoperation interfaces is to collect the remote information and improve visualization of the information to the operator so that they can comprehend in real-time and make informed decisions. We conducted formal user studies when possible and found that our approach and design can increase the operator's situation awareness

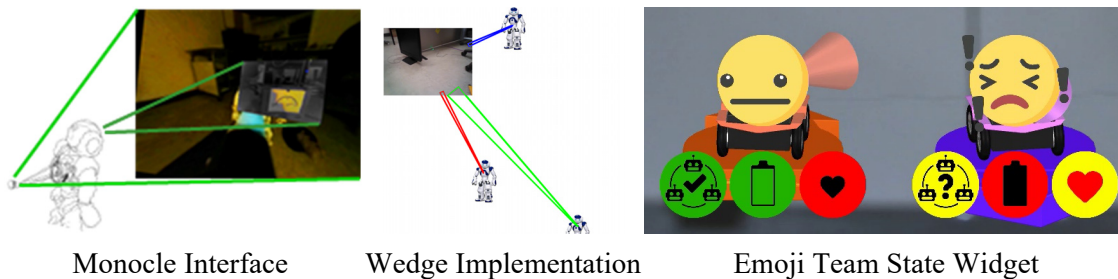


Figure 7-1. Examples of our design implementations from our projects.

helping them increase overall task performance. Our findings introduced in each chapter can help people in teleoperation fields, including search and rescue crews working in disaster zones, caregivers working with seniors, and other teleoperation enthusiasts.

In summary, our contributions are

- a novel interface design, prototype, and evaluation for simplifying interfaces that have multiple camera views into an environment;
- novel in-feed location indicators for team robot members to maintain team location awareness in a hectic team teleoperation scenario;
- design explorations and insights to present team robot members' working states through a series of user studies; and
- a novel interface design paradigm and its framework inspired by film and media techniques to provide teleoperation information during the control transition, helping the operator develop a high level of situation awareness.

We conducted the breadth-based exploration and presented them in four different chapters.

Our work leads to future research directions in improving teleoperation interface designs.

7.5. Toward Advanced Teleoperation Interfaces

For successful teleoperation, an operator must maintain significant awareness of robots,

remote environments, and mission tasks. Improving teleoperation interface designs can help the operator understand the necessary information within the limited time frame, directly related to their overall task performance. We presented this through this thesis.

Despite our breadth-based approach, we could not fully cover the teleoperation problem space. Our journey for improving teleoperation interface designs is yet to be done. Remaining assignments are including to find the balance between the detail and the abstraction of robot states, visualize multi-dimension information, present multi-modal sensor data, support the operator's projections of future states, design highly scalable interfaces for multi-sensor and multi-robot teleoperation, and so on.

Overall, our work assists to improve egocentric teleoperation interface designs by attacking pertinent problems facing teleoperation. Each project contributes to the more significant and ongoing discussions on teleoperation interface design elements. The discussions and results presented in this thesis will help interface designers and researchers guide the future development of teleoperation interfaces.

Bibliography

- Agarwal, R., Levinson, A. W., Allaf, M., Makarov, D. V., Nason, A., & Su, L.-M. (2007). The RoboConsultant: telementoring and remote presence in the operating room during minimally invasive urologic surgeries using a novel mobile robotic interface. *Urology*, 70(5), 970–974. <https://doi.org/10.1016/j.urology.2007.09.053>
- Al-Megren, S., & Ruddle, R. A. (2016). Comparing Tangible and Multi-touch Interaction for Interactive Data Visualization Tasks. *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, 279–286. <https://doi.org/10.1145/2839462.2839464>
- Bajracharya, M., Maimone, M. W., & Helmick, D. (2008). Autonomy for Mars Rovers: Past, present, and future. *Computer*, 41(12), 44–50. <https://doi.org/10.1109/MC.2008.479>
- Barber, D. J., Iv, J. A., Phillips, E., Talone, A. B., Jentsch, F., & Hill, S. G. (2015). Field Assessment of Multimodal Communication for Dismounted Human-Robot Teams. *59th Annual Meeting of the Human Factors and Ergonomics Society*, 921–925. <https://doi.org/10.1177/1541931215591280>
- Bark, K., Tran, C., Fujimura, K., & Ng-Thow-Hing, V. (2014). Personal Navi: Benefits of an Augmented Reality Navigational Aid Using a See-Thru 3D Volumetric HUD. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1–8). <https://doi.org/10.1145/2667317.2667329>
- Bartram, L., Ware, C., & Calvert, T. (2003). Moticons: detection, distraction and task. *International Journal of Human-Computer Studies*, 58(5), 515–545. [https://doi.org/10.1016/S1071-5819\(03\)00021-1](https://doi.org/10.1016/S1071-5819(03)00021-1)
- Baudisch, P., Good, N., Bellotti, V., & Schraedley, P. (2002). Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)* (pp. 259–266). Minneapolis, Minnesota, USA: ACM. <https://doi.org/10.1145/503376.503423>
- Baudisch, P., & Rosenholtz, R. (2003). Halo: a Technique for Visualizing Off-screen Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 481–488). Ft. Lauderdale, Florida, USA: ACM. <https://doi.org/10.1145/642611.642695>
- BBCClick. (2017). Fukushima disaster: The robots going where no human can. Retrieved May 13, 2020, from <https://www.bbc.com/news/av/technology-41584738/fukushima-disaster-the-robots-going-where-no-human-can>
- Bongshin Lee, Riche, N. H., Karlson, A. K., & Carpendale, S. (2010). SparkClouds: Visualizing Trends in Tag Clouds. *IEEE Transactions on Visualization and Computer Graphics*, 16(6), 1182–1189. <https://doi.org/10.1109/TVCG.2010.194>
- Borst, C., Suijkerbuijk, H. C. H., Mulder, M., & van Paassen, M. M. (2006). Ecological interface design for terrain awareness. *The International Journal of Aviation Psychology*, 16(4), 375–400.

- https://doi.org/10.1207/s15327108ijap1604_3
- Böttger, J., Preiser, M., Balzer, M., & Deussen, O. (2008). Detail-in-context visualization for satellite imagery. *Computer Graphics Forum*, 27(2), 587–596. <https://doi.org/10.1111/j.1467-8659.2008.01156.x>
- Brosz, J., Carpendale, S., & Nacenta, M. A. (2011). The Undistort Lens. *Computer Graphics Forum*, 30(3), 881–890. <https://doi.org/10.1111/j.1467-8659.2011.01937.x>
- Burigat, S., Chittaro, L., & Gabrielli, S. (2006). Visualizing Locations of Off-screen Objects on Mobile Devices: A Comparative Evaluation of Three Approaches. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services* (pp. 239–246). Helsinki, Finland: ACM. <https://doi.org/10.1145/1152215.1152266>
- Calhoun, G., Warfield, L., Wright, N., Spriggs, S., & Ruff, H. (2012). Automated Aid Evaluation for Transitioning UAS Camera Views. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 54(4), 413–417. <https://doi.org/10.1177/154193121005400430>
- Casper, J., & Murphy, R. R. (2003). Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 33(3), 367–385. <https://doi.org/10.1109/TSMCB.2003.811794>
- Chen, J., Glover, M., Li, C., & Yang, C. (2016). Development of a user experience enhanced teleoperation approach. In *2016 International Conference on Advanced Robotics and Mechatronics (ICARM)* (pp. 171–177). IEEE. <https://doi.org/10.1109/ICARM.2016.7606914>
- Chen, J. Y., & Barnes, M. J. (2013). *Human-Agent Teaming for Multi-Robot Control: A Literature Review*. Maryland. Retrieved from <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA583900>
- Chen, J. Y. C., & Barnes, M. J. (2012a). Supervisory Control of Multiple Robots: Effects of Imperfect Automation and Individual Differences. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54(2), 157–174. <https://doi.org/10.1177/0018720811435843>
- Chen, J. Y. C., & Barnes, M. J. (2012b). Supervisory control of multiple robots in dynamic tasking environments. *Ergonomics*, 55(9), 1043–1058. <https://doi.org/10.1080/00140139.2012.689013>
- Chen, J. Y. C., Barnes, M. J., & Qu, Z. (2010). RoboLeader: An agent for supervisory control of multiple robots, 81–82. <https://doi.org/10.1109/hri.2010.5453261>
- Chen, J. Y. C., Haas, E. C., & Barnes, M. J. (2007). Human Performance Issues and User Interface Design for Teleoperated Robots. *IEEE Transactions on Systems, Man and Cybernetics*, 37(6), 1231–1245. <https://doi.org/10.1109/TSMCC.2007.905819>
- Chien, S. Y., Lin, Y. L., Lee, P. J., Han, S., Lewis, M., & Sycara, K. (2018). Attention allocation for human multi-robot control: Cognitive analysis based on behavior data and hidden states. *International Journal of Human Computer Studies*, 000(June 2017), 1–15. <https://doi.org/10.1016/j.ijhcs.2018.03.005>
- Cohn, J. V., Older, B. A., Arnold, R. D., & Neill, E. B. O. (2013). Military Unmanned Aircraft System Operators : Training and Human Performance Issues, 341–346.
- Collins, C., Penn, G., & Carpendale, S. (2009). Bubble Sets: Revealing Set Relations with Isocontours over Existing Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 1009–1016. <https://doi.org/10.1109/TVCG.2009.122>
- Demir, M., McNeese, N. J., & Cooke, N. J. (2017). Team situation awareness within the context of human-autonomy teaming. *Cognitive Systems Research*, 46, 3–12. <https://doi.org/10.1016/j.cogsys.2016.11.003>
- Draper, M., Calhoun, G., Ruff, H., Mullins, B., Ayala, A., & Wright, N. (2008). Transition Display Aid for Changing Camera Views in UAV Operations. *Proceedings of the First Conference on Humans Operating Unmanned Systems (HUMOUS '08)*.
- Drury, J. L., Scholtz, J., & Yanco, H. a. H. A. (2003). Awareness in human-robot interactions. *IEEE International Conference on Systems, Man and Cybernetics*, 1(October), 912–918.

- <https://doi.org/10.1109/ICSMC.2003.1243931>
- Dubé, A. K., & McEwen, R. N. (2015). Do gestures matter? The implications of using touchscreen devices in mathematics instruction. *Learning and Instruction*, 40, 89–98. <https://doi.org/10.1016/j.learninstruc.2015.09.002>
- Durso, F. T., & Gronlund, S. D. (1999). Chapter 10: Situation Awareness. In *Handbook of Applied Cognition* (pp. 283–314). John Wiley & Sons, Inc.
- Dziubak, V., Dubois, P., Bunt, A., & Terry, M. (2016). Switter: supporting exploration of software learning materials on social media. *Proceedings of the ACM Conference on Designing Interactive Systems - DIS'16*, 1209–1220. <https://doi.org/10.1145/2901790.2901827>
- Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32–64. <https://doi.org/10.1518/001872095779049543>
- Endsley, M. R. (2011). Principles of Designing for SA. In *Designing for Situation Awareness: An Approach to User-Centered Design* (pp. 83–219). CRC Press. <https://doi.org/doi:10.1201/b11371-9>
- Endsley, M. R. (2015). Situation awareness: operationally necessary and scientifically grounded. *Cognition, Technology and Work*, 17(2), 163–167. <https://doi.org/10.1007/s10111-015-0323-5>
- Endsley, M. R. (2016). *Designing for Situation Awareness: An Approach to User-Centered Design* (Second Edi). CRC Press. Retrieved from <https://books.google.ca/books?id=eRPBkapAsggC>
- Ens, B. M., Finnegan, R., & Irani, P. P. (2014). The personal cockpit: a spatial interface for effective task switching on head-worn displays. In *ACM SIGCHI Human Factors in Computing Systems* (pp. 3171–3180). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2556288.2557058>
- Ferland, F., Pomerleau, F., Le Dinh, C. T., & Michaud, F. (2009). Egocentric and exocentric teleoperation interface using real-time, 3D video projection. *4th ACM/IEEE International Conference on Human Robot Interaction - HRI '09*, 37. <https://doi.org/10.1145/1514095.1514105>
- Freer, I., & Gibbs, O. (2018). Film Studies 101: The 30 Camera Shots Every Film Fan Needs To Know. Retrieved September 18, 2019, from <https://www.empireonline.com/movies/features/film-studies-101-camera-shots-styles/>
- Gittins, D. (1986). Icon-based human-computer interaction. *International Journal of Man-Machine Studies*, 24(6), 519–543. [https://doi.org/10.1016/S0020-7373\(86\)80007-4](https://doi.org/10.1016/S0020-7373(86)80007-4)
- Glas, D. F., Kanda, T., Ishiguro, H., & Hagita, N. (2012). Teleoperation of multiple social robots. In *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* (Vol. 42, pp. 530–544). IEEE. <https://doi.org/10.1109/TSMCA.2011.2164243>
- Gombolay, M., Bair, A., Huang, C., & Shah, J. (2017). Computational design of mixed-initiative human–robot teaming that considers human factors: situational awareness, workload, and workflow preferences. *The International Journal of Robotics Research*, 36(5–7), 597–617. <https://doi.org/10.1177/0278364916688255>
- Goodrich, M. A., & Schultz, A. C. (2007). Human-robot interaction: A survey. *Foundations and Trends in Human-Computer Interaction*, 1(3), 203–275. <https://doi.org/10.1561/11000000005>
- Gregori, N. Z., Lam, B. L., Gregori, G., & Ranganathan, S. (2012). Coherence Tomography in Patients and Carriers of X-linked Retinoschisis. *OPHTHA*, 120(1), 169–174. <https://doi.org/10.1016/j.ophtha.2012.07.051>
- Guo, C., Young, J. E., & Sharlin, E. (2009). Touch and toys: new techniques for interaction with a remote group of robots. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09* (p. 491). New York, New York, USA: ACM Press. <https://doi.org/10.1145/1518701.1518780>
- Gustafson, S. (2008). *Visualizing off-screen locations on small mobile displays*. Master's thesis. University of Manitoba. Retrieved from <http://gradworks.umi.com/MR/60/MR60483.html>
- Gustafson, S., Baudisch, P., Gutwin, C., & Irani, P. (2008). Wedge: Clutter-free Visualization of Off-screen

- Locations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 787–796). Florence, Italy: ACM. <https://doi.org/10.1145/1357054.1357179>
- Gustafson, S., & Irani, P. (2007). Comparing Visualizations for Tracking Off-screen Moving Targets. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems* (pp. 2399–2404). San Jose, CA, USA. <https://doi.org/10.1145/1240866.1241014>
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload* (Vol. 52, pp. 139–183). [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Hashimoto, S., Ishida, A., Inami, M., & Igarash, T. (2011). TouchMe: an augmented reality based remote robot manipulation. In *Proceedings of International Conference on Artificial Reality and Telexistence (ICAT)* (pp. 1–6).
- Hayes, B., & Shah, J. A. (2017). Improving Robot Controller Transparency Through Autonomous Policy Explanation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI '17* (pp. 303–312). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2909824.3020233>
- Herring, S. C., Fussell, S. R., Kristoffersson, A., Mutlu, B., Neustaedter, C., & Tsui, K. (2016). The Future of Robotic Telepresence. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16* (pp. 1038–1042). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2851581.2886423>
- Hess, U., & Hareli, S. (2015). The Role of Social Context for the Interpretation of Emotional Facial Expressions. In M. K. Mandal & A. Awasthi (Eds.), *Understanding Facial Expressions in Communication* (pp. 119–141). New Delhi: Springer India. https://doi.org/10.1007/978-81-322-1934-7_7
- Hewett, T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., ... Verplank, W. (1992). *ACM SIGCHI Curricula for Human-Computer Interaction. Journal of Chemical Information and Modeling* (Vol. 53). Association for Computing Machinery. <https://doi.org/10.1145/2594128>
- Hill, B. (2016). *The Magic of Fiction: Crafting Words into Story: The Writer's Guide to Writing & Editing*. Title Page Books; Second edition.
- Hong, A., Lee, D. G., Bülthoff, H. H., & Son, H. Il. (2016). Multimodal Feedback for Teleoperation of Multiple Mobile Robots in an Outdoor Environment. *Journal on Multimodal User Interfaces*, 1–14. <https://doi.org/10.1007/s12193-016-0230-y>
- Htun, N. N., Halvey, M., & Baillie, L. (2017). An Interface for Supporting Asynchronous Multi-Level Collaborative Information Retrieval. *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval - CHIIR '17*, 225–234. <https://doi.org/10.1145/3020165.3020172>
- Hughes, S., & Lewis, M. (2004). Robotic camera control for remote exploration. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 6(1), 511–517. <https://doi.org/10.1145/985692.985757>
- Ion, A., Chang, Y.-L. B., Haller, M., Hancock, M., & Scott, S. D. (2013). Canyon: Providing Location Awareness of Multiple Moving Objects in a Detail View on Large Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3149–3158). Paris, France: ACM.
- Jaju, A., Banerji, A., & Pal, P. K. (2013). Development and evaluation of a telepresence interface for teleoperation of a robot manipulator. In *2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)* (pp. 90–95). IEEE. <https://doi.org/10.1109/URAI.2013.6677481>
- Jia, Y., Xu, B., Shen, J., Pei, M., Dong, Z., Hou, J., & Yang, M. (2015). Telepresence Interaction by Touching Live Video Images. *Human-Computer Interaction*. Retrieved from <http://arxiv.org/abs/1512.04334>
- Jiang, X., Zheng, B., Bednarik, R., & Atkins, M. S. (2015). Pupil responses to continuous aiming movements. *International Journal of Human Computer Studies*, 83, 1–11. <https://doi.org/10.1016/j.ijhcs.2015.05.006>

- Julio, R. E., & Bastos, G. S. (2015). Dynamic bandwidth management library for multi-robot systems. *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2585–2590. <https://doi.org/10.1109/IROS.2015.7353729>
- Kamaruddin, N., & Sulaiman, S. (2018). Understanding Interface Design Principles and Elements Guidelines: A Content Analysis of Established Scholars. In *Proceedings of the Art and Design International Conference (AnDIC 2016)* (pp. 89–100). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-0487-3_11
- Katz, S. D. (2019). *Film Directing: Shot by Shot: Visualizing from Concept to Screen*. Michael Wiese Productions; Anniversary edition.
- Keahey, T. A. (1998). The generalized detail in-context problem. In *Proceedings of the IEEE Symposium on Information Visualization* (pp. 44–51). IEEE. <https://doi.org/10.1109/INFVIS.1998.729558>
- Kelly, A., Chan, N., Herman, H., Huber, D., Meyers, R., Rander, P., ... Capstick, E. (2010). Real-time photorealistic virtualized reality interface for remote mobile robot control. *The International Journal of Robotics Research*, 30(3), 384–404. <https://doi.org/10.1177/0278364910383724>
- Khan, A., Fitzmaurice, G., Almeida, D., Burtnyk, N., & Kurtenbach, G. (2004). A Remote Control Interface for Large Displays. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology* (pp. 127–136). Santa Fe, NM, USA: ACM. <https://doi.org/10.1145/1029632.1029655>
- Kidwell, B., Calhoun, G. L., Ruff, H. A., & Parasuraman, R. (2012). Adaptable and adaptive automation for supervisory control of multiple autonomous vehicles. *Proceedings of the Human Factors and Ergonomics Society*, (2007), 428–432. <https://doi.org/10.1177/1071181312561096>
- Kim, S., & Dey, A. K. (2009). Simulated Augmented Reality Windshield Display As a Cognitive Mapping Aid for Elder Driver Navigation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 133–142). Boston, MA, USA: ACM. <https://doi.org/10.1145/1518701.1518724>
- Kline, T. J. B., Ghali, L. M., Kline, D. W., & Brown, S. (1990). Visibility Distance of Highway Signs among Young, Middle-Aged, and Older Observers: Icons Are Better than Text. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 32(5), 609–619. <https://doi.org/10.1177/001872089003200508>
- Kobsa, A., Koenemann, J., & Pohl, W. (2001). Personalised hypermedia presentation techniques for improving online customer relationships. *The Knowledge Engineering Review*, 16(02), 111. <https://doi.org/10.1017/S0269888901000108>
- Kolling, A., Walker, P., Chakraborty, N., Sycara, K., & Lewis, M. (2016). Human Interaction With Robot Swarms: A Survey. *IEEE Transactions on Human-Machine Systems*, 46(1), 9–26. <https://doi.org/10.1109/THMS.2015.2480801>
- Komiyama, R., Miyaki, T., & Rekimoto, J. (2017). JackIn Space: Designing a Seamless Transition Between First and Third Person View for Effective Telepresence Collaborations. In *Proceedings of the 8th Augmented Human International Conference on - AH '17* (pp. 1–9). New York, New York, USA: ACM Press. <https://doi.org/10.1145/3041164.3041183>
- Kruckel, K., Nolden, F., Ferrein, A., & Scholl, I. (2015). Intuitive visual teleoperation for UGVs using free-look augmented reality displays. *IEEE International Conference on Robotics and Automation (ICRA)*, 4412–4417. <https://doi.org/10.1109/ICRA.2015.7139809>
- Labonte, D., Boissy, P., & Michaud, F. (2010). Comparative analysis of 3-D robot teleoperation interfaces with novice users. *IEEE International Conference on Systems, Man and Cybernetics*, 40(5), 1331–1342. <https://doi.org/10.1109/TSMCB.2009.2038357>
- Lee, D., Franchi, A., Son, H. Il, Ha, C., Bulthoff, H. H., & Giordano, P. R. (2013). Semiautonomous Haptic Teleoperation Control Architecture of Multiple Unmanned Aerial Vehicles. *IEEE/ASME Transactions on Mechatronics*, 18(4), 1334–1345. <https://doi.org/10.1109/TMECH.2013.2263963>
- Lee, L., Romano, R., & Stein, G. (2000). Monitoring activities from multiple video streams: Establishing a common coordinate frame. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8),

- 758–767. <https://doi.org/10.1109/34.868678>
- Leeper, A. E., Hsiao, K., Ciocarlie, M., Takayama, L., & Gossow, D. (2012). Strategies for human-in-the-loop robotic grasping. In *Proceedings of the 7th annual ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 1–8). <https://doi.org/10.1145/2157689.2157691>
- Lewis, M. (2013). Human Interaction With Multiple Remote Robots. *Reviews of Human Factors and Ergonomics*, 9(1), 131–174. <https://doi.org/10.1177/1557234X13506688>
- Li, X. (2014). Research on the Montage Technique in The Film and Television, (Iceland), 599–603.
- Liu, Y., & Nejat, G. (2013). Robotic urban search and rescue: A survey from the control perspective. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 72(2), 147–165. <https://doi.org/10.1007/s10846-013-9822-x>
- Long, G. M., & Kearns, D. F. (1996). Visibility of Text and Icon Highway Signs under Dynamic Viewing Conditions. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 38(4), 690–701. <https://doi.org/10.1518/001872096778827215>
- Mao, Y., Wang, W., Ding, C., Guo, W., Jiang, X., Baumann, M., & Wets, G. (2018). A measurement to driving situation awareness in signalized intersections. *Transportation Research Part D: Transport and Environment*, 62(July 2017), 739–747. <https://doi.org/10.1016/j.trd.2018.05.001>
- Mast, M., Materna, Z., Španěl, M., Weisshardt, F., Arbeiter, G., Burmester, M., ... Graf, B. (2014). Semi-autonomous domestic service robots: evaluation of a user interface for remote manipulation and navigation with focus on effects of stereoscopic display. *International Journal of Social Robotics*, 7(2), 183–202. <https://doi.org/10.1007/s12369-014-0266-7>
- Meghdadi, A. H., & Irani, P. (2013). Interactive Exploration of Surveillance Video through Action Shot Summarization and Trajectory Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), 2119–2128. <https://doi.org/10.1109/TVCG.2013.168>
- Murch, G. M. (1985). Using color effectively: Designing to human specifications. *Technical Communication*, 32(4), 14–20. Retrieved from <http://www.jstor.org/stable/43094557>
- Nielsen, C. W., Goodrich, M. A., & Ricks, R. W. (2007). Ecological interfaces for improving mobile robot teleoperation. *IEEE Transactions on Robotics*, 23(5), 927–941. <https://doi.org/10.1109/TRO.2007.907479>
- O’Keeffe, S., Ward, T. E., & Villing, R. (2016). Improving Task Performance through High Level Shared Control of Multiple Robots with a Context Aware Human-Robot Interface. In *2016 International Conference on Autonomous Robot Systems and Competitions (ICARSC)* (pp. 277–282). IEEE. <https://doi.org/10.1109/ICARSC.2016.45>
- Ochiai, Y., Takemura, K., Ikeda, A., Takamatsu, J., & Ogasawara, T. (2014). Remote control system for multiple mobile robots using touch panel interface and autonomous mobility. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Iros), 3272–3277. <https://doi.org/10.1109/IROS.2014.6943017>
- Odegaard Writing & Research Center. (n.d.). Using Transitions Effectively. Retrieved July 16, 2019, from [http://depts.washington.edu/owrc/Handouts/Using Transitions Effectively.pdf](http://depts.washington.edu/owrc/Handouts/Using%20Transitions%20Effectively.pdf)
- Omidshafiei, S., Agha–Mohammadi, A., Amato, C., Liu, S., How, J. P., & Vian, J. (2017). Decentralized control of multi-robot partially observable Markov decision processes using belief space macro-actions. *The International Journal of Robotics Research*, 36(2), 231–258. <https://doi.org/10.1177/0278364917692864>
- Parasuraman, R., Caccamo, S., Fredrik, B., Ogren, P., & Neerincx, M. (2017). A New UGV Teleoperation Interface for Improved Awareness of Network Connectivity and Physical Surroundings, 6(3), 48–70. <https://doi.org/10.5898/JHRI.6.3.Parasuraman>
- Parkinson, B. (2005). Do Facial Movements Express Emotions or Communicate Motives? *Personality and Social Psychology Review*, 9(4), 278–311. https://doi.org/10.1207/s15327957pspr0904_1
- Paul, C., & Komlodi, A. (2012). Emotion as an indicator for future interruptive notification experiences. In

- Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12* (p. 2003). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2212776.2223743>
- Peskoe-Yang, L. (2019). Paris Firefighters Used This Remote-Controlled Robot to Extinguish the Notre Dame Blaze. Retrieved May 13, 2020, from <https://spectrum.ieee.org/automaton/robotics/industrial-robots/colossus-the-firefighting-robot-that-helped-save-notre-dame>
- Phillips, E. K., & Jentsch, F. G. (2017). Supporting situation awareness through robot-to-human information exchanges under conditions of visuospatial perspective taking. *Journal of Human-Robot Interaction*, 6(3), 92–117. <https://doi.org/10.5898/JHRI.6.3.Phillips>
- Pingali, G., Opalach, A., Jean, Y., & Carlbom, I. (2001). Visualization of Sports Using Motion Trajectories: Providing Insights into Performance, Style, and Strategy. *Proceedings of the Conference on Visualization '01*, 75–82. <https://doi.org/10.1109/VISUAL.2001.964496>
- Potter, K., Wilson, A., Bremer, P. T., Williams, D., Doutriaux, C., Pascucci, V., & Johnson, C. R. (2009). Ensemble-vis: A framework for the statistical visualization of ensemble data. *ICDM Workshops 2009 - IEEE International Conference on Data Mining*, 233–240. <https://doi.org/10.1109/ICDMW.2009.55>
- Price, T. F., & LaFiandra, M. (2017). The perception of team engagement reduces stress induced situation awareness overconfidence and risk-taking. *Cognitive Systems Research*, 46, 52–60. <https://doi.org/10.1016/j.cogsys.2017.02.004>
- Rea, D. J., Seo, S. H., Bruce, N., & Young, J. E. (2017). Movers, Shakers, and Those Who Stand Still: Visual Attention-grabbing Techniques in Robot Teleoperation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI '17* (pp. 398–407). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2909824.3020246>
- Rea, D. J., Seo, S. H., & Young, J. E. (2020). Social Robotics for Nonsocial Teleoperation: Leveraging Social Techniques to Impact Teleoperator Performance and Experience. *Current Robotics Reports*, 1(4), 287–295. <https://doi.org/10.1007/s43154-020-00020-7>
- Rea, D. J., & Young, J. E. (2018). It's All in Your Head: using priming to shape an operator's perceptions and behavior during teleoperation, 32–40. <https://doi.org/10.1145/3171221.3171259>
- Recchiuto, C. T., Sgorbissa, A., & Zaccaria, R. (2016). Visual Feedback with Multiple Cameras in a UAVs Human-Swarm Interface. *Robotics and Autonomous Systems*, 80, 43–54. <https://doi.org/10.1016/j.robot.2016.03.006>
- Ribeiro, D., Bernardes, J., Roman, N., Antunes, M., Ortega, E., Sousa, A., ... Lima, C. (2016). Analysis of choreographed human movements using depth cameras: a systematic review. In *Human-Computer Interaction. Interaction Platforms and Techniques* (pp. 82–92). Toronto, ON, Canada: Springer International Publishing. https://doi.org/10.1007/978-3-319-39516-6_8
- Richer, J., & Drury, J. L. (2006). A Video Game-based Framework for Analyzing Human-robot Interaction: Characterizing Interface Design in Real-time Interactive Multimedia Applications. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction* (pp. 266–273). Salt Lake City, Utah, USA: ACM. <https://doi.org/10.1145/1121241.1121287>
- Riek, L. (2012). Wizard of Oz studies in HRI: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 1(1), 119–136. <https://doi.org/10.5898/JHRI.1.1.Riek>
- Rogers, Y., Sharp, H., & Preece, J. (2007). *Interaction Design: beyond human-computer interaction* (2nd ed.). John Wiley & Sons.
- Roldán, J. J., Peña-Tapia, E., Martín-Barrio, A., Olivares-Méndez, M. A., del Cerro, J., & Barrientos, A. (2017). Multi-robot interfaces and operator situational awareness: Study of the impact of immersion and prediction. *Sensors (Switzerland)*, 17(8), 1–25. <https://doi.org/10.3390/s17081720>
- Rosenfeld, A., Agmon, N., Maksimov, O., & Kraus, S. (2017). Intelligent agent supporting human-multi-robot team collaboration. *Artificial Intelligence*, 252, 211–231. <https://doi.org/10.1016/j.artint.2017.08.005>
- Rosenfeld, A., Noa, A., Maksimov, O., & Kraus, S. (2016). Human-multi-robot team collaboration for

- efficient warehouse operation. *Autonomous Robots and Multirobot Systems (ARMS)*.
- Saakes, D., Choudhary, V., Sakamoto, D., Inami, M., Lgarashi, T., & Igarashi, T. (2013). A teleoperating interface for ground vehicles using autonomous flying cameras. In *International Conference on Artificial Reality and Telexistence (ICAT)* (pp. 13–19). IEEE. <https://doi.org/10.1109/ICAT.2013.6728900>
- Saffer, D. (2013). Chapter 4: Feedback. In *Microinteractions: designing with details* (pp. 83–107). O'Reilly Media, Inc.
- Sakamoto, D., Honda, K., Inami, M., & Igarashi, T. (2009). Sketch and run: a stroke-based interface for home robots. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)* (pp. 197–200). <https://doi.org/10.1145/1518701.1518733>
- Salmanipour, S., & Sirouspour, S. (2013). Teleoperation of a mobile robot with model-predictive obstacle avoidance control. In *Annual Conference of the IEEE Industrial Electronics Society (IECON)* (pp. 4270–4275). <https://doi.org/10.1109/IECON.2013.6699821>
- Selkowitz, A. R., Lakhmani, S. G., & Chen, J. Y. C. (2017). Using agent transparency to support situation awareness of the Autonomous Squad Member. *Cognitive Systems Research*, 46, 13–25. <https://doi.org/10.1016/j.cogsys.2017.02.003>
- Seo, S. H., Geiskovitch, D., Nakane, M., King, C., & Young, J. E. (2015). Poor Thing! Would You Feel Sorry for a Simulated Robot? In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '15* (pp. 125–132). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2696454.2696471>
- Seo, S. H., Rea, D. J., Wiebe, J., & Young, J. E. (2017). Monocle: Interactive detail-in-context using two pan-and-tilt cameras to improve teleoperation effectiveness. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 962–967). Lisbon, Portugal: IEEE. <https://doi.org/10.1109/ROMAN.2017.8172419>
- Seo, S. H., Takashima, K., Young, J. E., & Kitamura, Y. (2014). Alerting Users by Animating Content on a Transforming Tabletop Interface. In *Human Interface Symposium*. Kyoto, Japan.
- Seo, S. H., Young, J. E., & Irani, P. (2017). Where are the robots? In-feed embedded techniques for visualizing robot team member locations. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 522–527). Lisbon, Portugal: IEEE. <https://doi.org/10.1109/ROMAN.2017.8172352>
- Sharma, M., Hildebrandt, D., Newman, G., Young, J. E., & Eskicioglu, R. (2013). Communicating affect via flight path Exploring use of the Laban Effort System for designing affective locomotion paths. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 293–300). IEEE. <https://doi.org/10.1109/HRI.2013.6483602>
- Singh, A., Seo, S. H., Hashish, Y., Nakane, M., Young, J. E., & Bunt, A. (2013). An interface for remote robotic manipulator control that reduces task load and fatigue. In *2013 IEEE RO-MAN* (pp. 738–743). IEEE. <https://doi.org/10.1109/ROMAN.2013.6628401>
- Singh, A., & Young, J. (2013). A Dog Tail for Utility Robots: exploring affective properties of tail movement. *Human-Computer Interaction—INTERACT 2013*, 8118, 403–419. https://doi.org/10.1007/978-3-642-40480-1_27
- Son, H. Il, Franchi, A., Chuang, L. L., Kim, J., Bülthoff, H. H., & Robuffo Giordano, P. (2013). Human-Centered Design and Evaluation of Haptic Cueing for Teleoperation of Multiple Mobile Robots. *IEEE Transactions on Cybernetics*, 43(2), 597–609. <https://doi.org/10.1109/TSMCB.2012.2212884>
- Squire, P. N., & Parasuraman, R. (2010). Effects of automation and task load on task switching during human supervision of multiple semi-autonomous robots in a dynamic environment. *Ergonomics*, 53(8), 951–961. <https://doi.org/10.1080/00140139.2010.489969>
- Sternberg, S. (1969). Memory-scanning: Mental Processes Revealed by Reaction-time Experiments. *American Scientist*, 57(4), 421–457.
- Student Writing Support. (n.d.). Sentence transitions. Retrieved September 24, 2019, from

<http://writing.umn.edu/sws/quickhelp/style/transitions.html>

- Tesla, N. (1898). *613809*. United States. Retrieved from <https://www.google.com/patents/US613809>
- Tonniss, M., & Klinker, G. (2006). Effective Control of a Car Driver's Attention for Visual and Acoustic Guidance Towards the Direction of Imminent Dangers. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality* (pp. 13–22). IEEE Computer Society. <https://doi.org/10.1109/ISMAR.2006.297789>
- Uyttendaele, M., Criminisi, A., Kang, S. B., Winder, S., Szeliski, R., & Hartley, R. (2004). Image-based interactive exploration of real-world environments. *IEEE Computer Graphics and Applications*, 24(3), 52–63. <https://doi.org/10.1109/MCG.2004.1297011>
- Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, 41(2), 167–174. <https://doi.org/10.1111/j.1469-8986.2003.00148.x>
- Wang Baldonado, M. Q., Woodruff, A., & Kuchinsky, A. (2000). Guidelines for using multiple views in information visualization. In *Proceedings of the working conference on Advanced visual interfaces - AVI '00* (Vol. 26, pp. 110–119). New York, New York, USA: ACM Press. <https://doi.org/10.1145/345513.345271>
- Wentzel, J., Rea, D. J., Young, J. E., & Sharlin, E. (2015). Shared presence and collaboration using a co-located humanoid robot. *HAI 2015 - Proceedings of the 3rd International Conference on Human-Agent Interaction*, 273–276. <https://doi.org/10.1145/2814940.2814995>
- Wong, C. Y., & Seet, G. (2017). Workload, awareness and automation in multiple-robot supervision. *International Journal of Advanced Robotic Systems*, 14(3), 1–16. <https://doi.org/10.1177/1729881417710463>
- Wong, N., Carpendale, S., & Greenberg, S. (2005). Edgelens: an interactive method for managing edge congestion in graphs. In *IEEE Symposium on Information Visualization 2003 (IEEE Cat. No.03TH8714)* (Vol. 2003, pp. 51–58). IEEE. <https://doi.org/10.1109/INFVIS.2003.1249008>
- Wright, J. L., Chen, J. Y. C., & Barnes, M. J. (2018). Human–automation interaction for multiple robot control: the effect of varying automation assistance and individual differences on operator performance. *Ergonomics*, 61(8), 1033–1045. <https://doi.org/10.1080/00140139.2018.1441449>
- Yanco, H. A., & Drury, J. (2004a). Classifying human-robot interaction: an updated taxonomy. *IEEE Systems, Man and Cybernetics*, 3, 2841–2846. <https://doi.org/10.1109/ICSMC.2004.1400763>
- Yanco, H. A., & Drury, J. (2004b). “Where am I?” Acquiring situation awareness using a remote robot platform. In *IEEE International Conference on Systems, Man and Cybernetics* (Vol. 3, pp. 2835–2840). <https://doi.org/10.1109/ICSMC.2004.1400762>
- Yang, J., Kamezaki, M., Sato, R., Iwata, H., & Sugano, S. (2015). Inducement of visual attention using augmented reality for multi-display systems in advanced tele-operation. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 5364–5369). IEEE. <https://doi.org/10.1109/IROS.2015.7354135>
- Zhang, F., & Liu, F. (2014). Parallax-tolerant image stitching. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3262–3269). <https://doi.org/10.1109/CVPR.2014.423>
- Zheng, K., Glas, D. F., Kanda, T., & Ishiguro, H. (2012). Designing and implementing a human-robot team for social interactions. *Systems, Man, and Cybernetics: Systems*, 43(4), 1–17.
- Zheng, K., Glas, D. F., Kanda, T., Ishiguro, H., & Hagita, N. (2013). Supervisory Control of Multiple Social Robots for Navigation. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction* (Vol. 3, pp. 17–24). IEEE Press. <https://doi.org/10.1109/HRI.2013.6483497>

Appendix A Multiple Cameras: Interactive Detail-in-context Interface Using Multiple Cameras in Teleoperation

- Research Ethics and Compliance Approval Certificate
- Public Recruitment Poster
- Consent Form
- Demographic Questionnaire
- Post-condition Questionnaire
- Post-study Questionnaire
- Semi-structured Interview Protocol
- Meter Report Sheet
- Meters

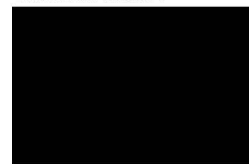


UNIVERSITY
OF MANITOBA

Research Ethics and Compliance

Office of the Vice-President (Research and International)

Human Ethics



APPROVAL CERTIFICATE

April 1, 2015

NSERC

TO: James E. Young
Principal Investigator [REDACTED]

FROM: Susan Frohlick, Chair
Joint-Faculty Research Ethics Board (JFREB)

Re: Protocol #J2015:020
"Exploring alternate camera systems for remotely controlling robots"

Please be advised that your above-referenced protocol has received human ethics approval by the **Joint-Faculty Research Ethics Board**, which is organized and operates according to the Tri-Council Policy Statement (2). **This approval is valid for one year only.**

Any significant changes of the protocol and/or informed consent form should be reported to the Human Ethics Secretariat in advance of implementation of such changes.

Please note:

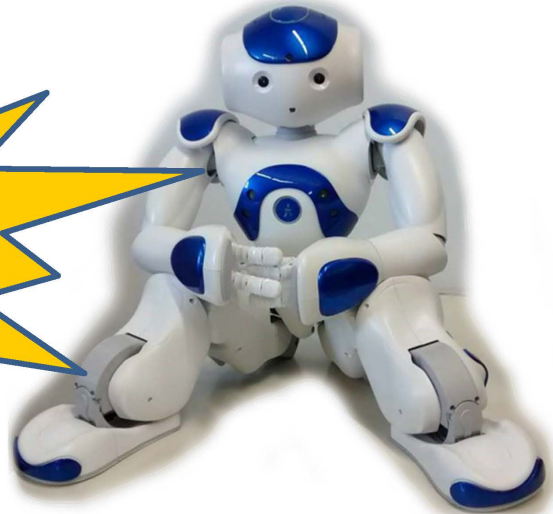
- If you have funds pending human ethics approval, please mail/e-mail/fax ([REDACTED]) a copy of this Approval (identifying the related UM Project Number) to the Research Grants Officer in ORS in order to initiate fund setup. (How to find your UM Project Number: <http://umanitoba.ca/research/ors/mrt-faq.html#pr0>)
- if you have received multi-year funding for this research, responsibility lies with you to apply for and obtain Renewal Approval at the expiry of the initial one-year approval; otherwise the account will be locked.

The Research Quality Management Office may request to review research documentation from this project to demonstrate compliance with this approved protocol and the University of Manitoba *Ethics of Research Involving Humans*.

The Research Ethics Board requests a final report for your study (available at: http://umanitoba.ca/research/orec/ethics/human_ethics_REB_forms_guidelines.html) in order to be in compliance with Tri-Council Guidelines.



UNIVERSITY
OF MANITOBA



Control NAO (a Humanoid Robot) as you complete various tasks in a one-hour human-robot interaction experiment at the University of Manitoba. Note that you must be over 18 to participate in our experiment.

Please visit:

[http://\[redacted\]](http://[redacted]) or

[http://\[redacted\]](http://[redacted])

If you have any questions about the study, please contact Stela H. Seo at [redacted] or Dr. James E. Young at [redacted].

This research study was approved by the Joint-Faculty Research Ethics Board, University of Manitoba

Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo





Project Title: The Dynamic Third-Person View Teleoperation Interface

Researchers: Dr. James E. Young, Stela H. Seo, Daniel J. Rea



This consent form, a copy of which will be left with you for your records and reference, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

Participation in this study is voluntary, and will take approximately 60 minutes of your time. For this study, you will have the opportunity to control a humanoid robot (called NAO) to simulate aspects of the discovery of victims and inspecting valve meters in a search and rescue mission. NAO is a 58 cm tall humanoid (walking) robot. You will use live camera feeds to remotely control the NAO robot. You will view the video and control the robot via a computer. This study will include 2 main phases. To begin, you will observe NAO in its simulated search and rescue environment. You will learn to control NAO and observe its interaction in the environment. For the main portion of the experiment you will be brought to a second room within the same building, visually removed from NAO and its environment. At this point you will be given a number of tasks to assess the suitability of various interfaces for remotely controlling NAO during various tasks. No expertise or experience is necessary. You will receive \$10 for your participation.

All information you provide is considered completely confidential; your name will not be included, or in any other way associated, with the data collected in the study. Data collected during this study will be used for academic research and purpose of publication in an anonymous form. We may use anonymized video or audio data for purposes of public presentation and dissemination only with your express permission (given below). In addition, data will be retained for a maximum of five years in a locked office in the EITC building, University of Manitoba, to which only researchers associated with this study have access. Once published, results of the study will be made available to the public for free at <http://home.cs.umanitoba.ca/~young/>. Again, no personal information about your involvement will be included. Please note that the University of Manitoba may look at the research records to see that the research is being done in a safe and proper way.

Your signature on this form indicates that you have understood, to your satisfaction, the information regarding participation in the research project and agree to participate as a subject. By doing this you also confirm that you are of the age of majority in Canada (18 years or more). In no way does this waive your legal rights nor release the researchers, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from the study at any time, and to refrain from answering any questions asked, without prejudice or consequence. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

This research has been approved by the Joint-Faculty Research Ethics Board. If you have any concerns or complaints about this project you may contact Dr. James Young at [redacted] or the Human Ethics Secretariat at [redacted]. A copy of this consent form has been given to you to keep for your records and reference.

For purposes of research and analysis it is necessary for the experiment to be videotaped.

Do you agree that any video footage taken may also be used for distribution of research, for example, through research videos or images taken from your video? If you say No, your video will be used for internal data analysis purposes only.

No ___ Yes ___ but only if you blur my face ___ AND/OR if you muffle my voice ___

Participant's Name _____ Signature _____ Date _____

Researcher's Name _____ Signature _____ Date _____

Demographics Questionnaire

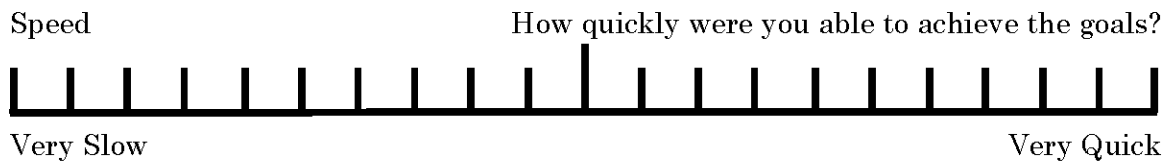
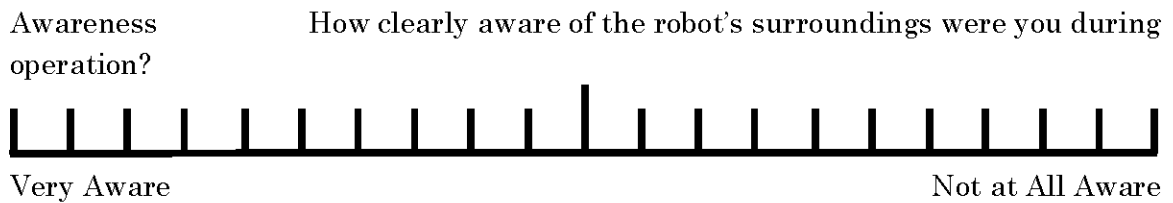
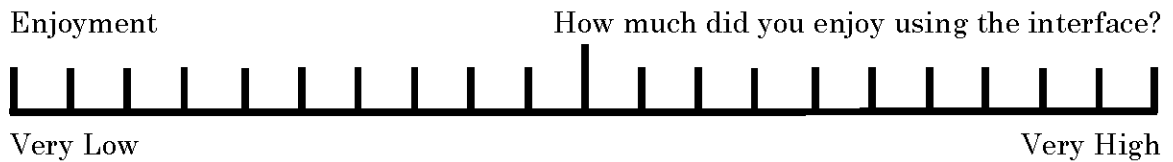
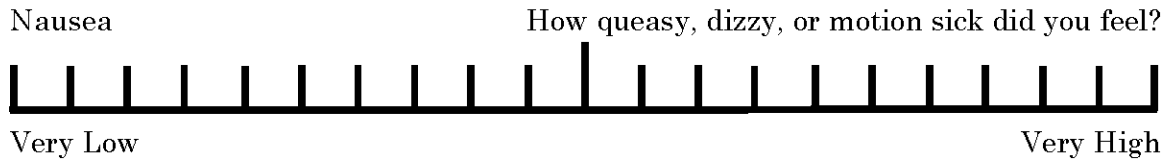
PARTICIPANT ID: ____

- 1) What is your age?
18-20____ 21-25____ 26-30____ 31-35____ 36-40____ 40+____
- 2) What is your sex?
Male____ Female____ Intersex____
- 3) How often do you play 3D videogames such as shooters, racing games...?
____Never played videogames
____A few times a month or less
____Once a week
____More than once a week
- 4) How would you rate your current skill level for this kind of 3D video game?
Very poor____ Poor____ Fair____ Good____ Very good____
- 5) How often do you drive a motor vehicle (car, motorcycle, etc.)?
____Never driven a vehicle
____A few times a month or less
____Once a week
____More than once a week
- 6) How would you rate your current vehicle driving skill level?
Very poor____ Poor____ Fair____ Good____ Very good____
- 7) How often do you remotely control a vehicle (e.g. car, plane, drone, quadcopter, robot, etc.)?
____Never remotely controlled a vehicle
____A few times
____Every few months
____Several times a month

Post-Condition Questionnaire

PARTICIPANT ID: ____

For the following question, mark **ONE** position along the scale:



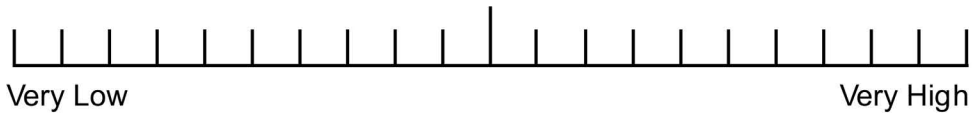
Post-Condition Questionnaire

PARTICIPANT ID: _____

For the following question, mark **ONE** position along the scale:

Mental Demand

How mentally demanding was the task?



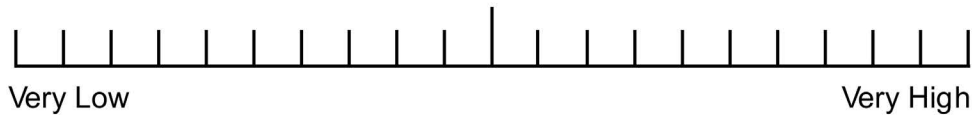
Physical Demand

How physically demanding was the task?



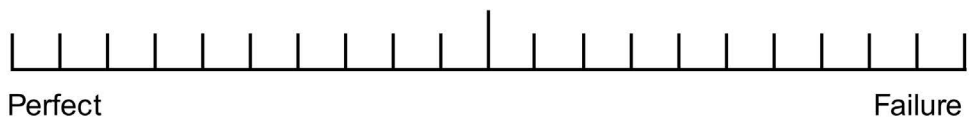
Temporal Demand

How hurried or rushed was the pace of the task?



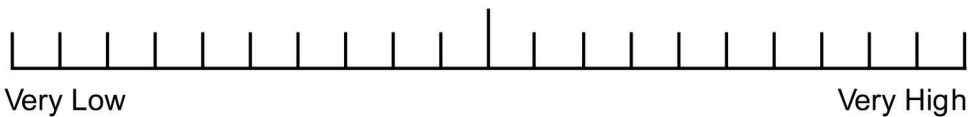
Performance

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



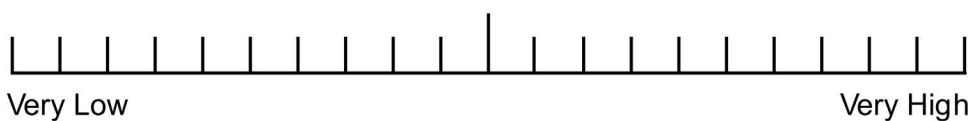
Effort

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



Frustration

How insecure, discouraged, irritated, stressed, and annoyed were you?



General Experience Questionnaire

PARTICIPANT ID: ____

- 1) A) Please rank the interfaces from **1** (being the interface that you **MOST** preferred) to **3** (being the interface that you **LEAST** preferred).

____ Toggle View Interface
____ Side by Side Interface
____ Contextual Merged Interface

B) Please describe any pros and cons that you found with each interface.

Toggle View Interface

Pros:

Cons: _____

Side by Side Interface

Pros:

Cons: _____

Contextual Merged Interface

Pros:

Cons: _____

- 2) Did you experience any motion sickness at all? What do you think may have caused or contributed to this?

3) Do you have any additional positive comments?

Do you have any additional negative comments?

Do you have any final comments or suggestions?

Semi-Structured Interview Protocol:

The semi-structured interview will focus on participants leading the discussion. We use the following as the guiding principles of the interview.

Goals and Benefits:

Primary research goal: Discover insight into how the differences between interfaces affect a participant's ability to perform tasks, and how the interfaces may be well or poorly suited to the tasks we selected.

Secondary research goal: Elicit feedback on potential future improvements to our interfaces

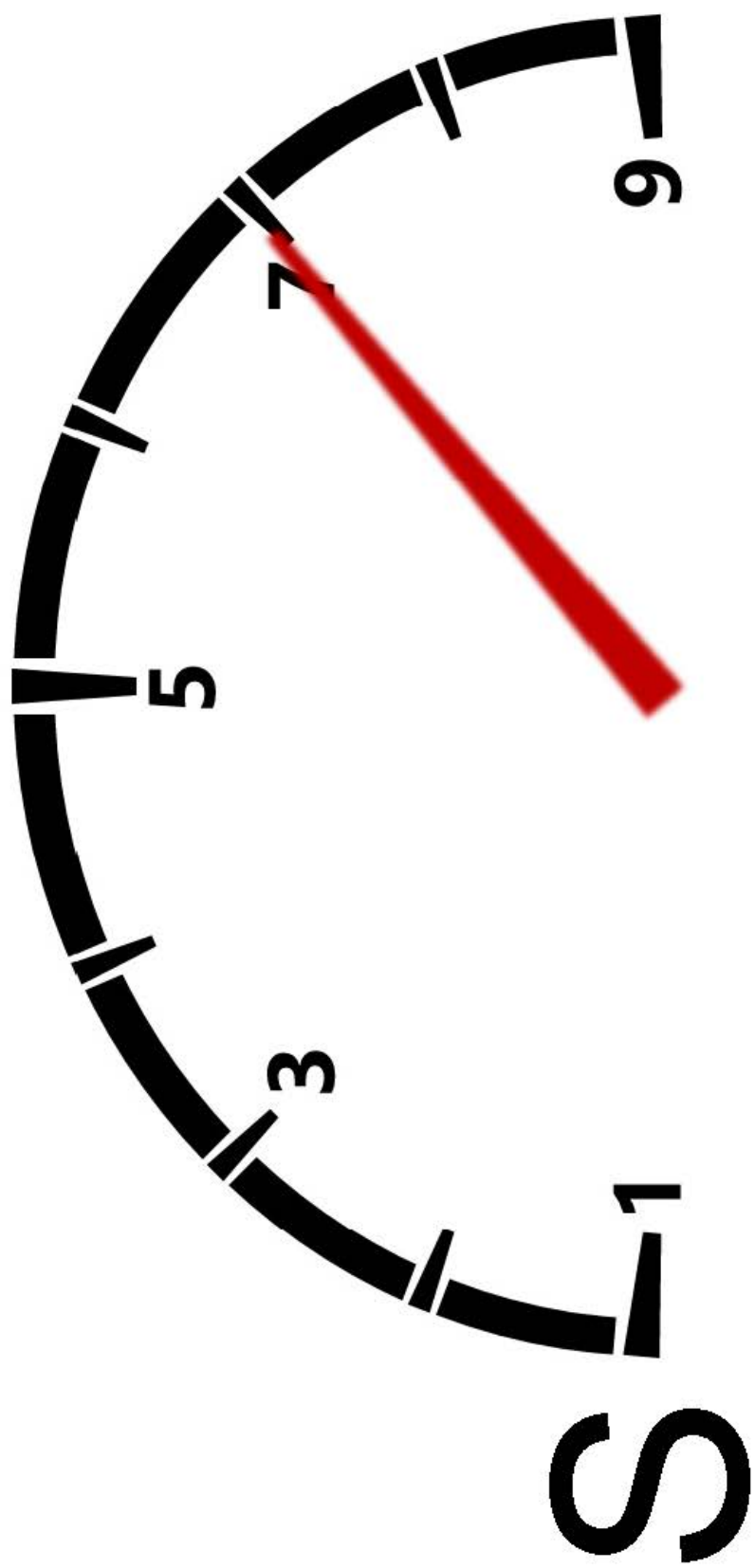
Benefits to participant: Provide them with a detailed opportunity to inquire further about the work and to enable us to answer any questions or concerns that they may have.

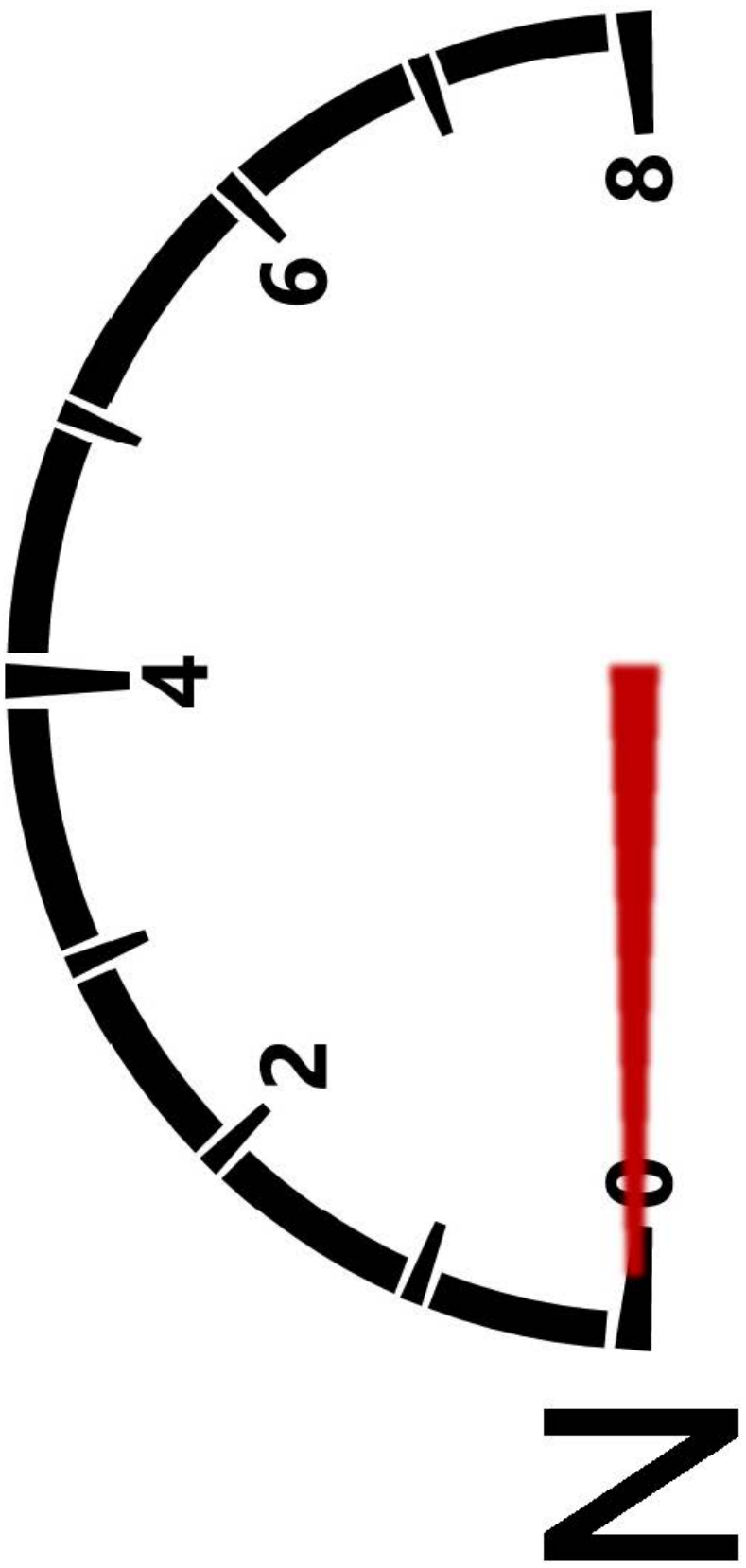
Guiding Questions:

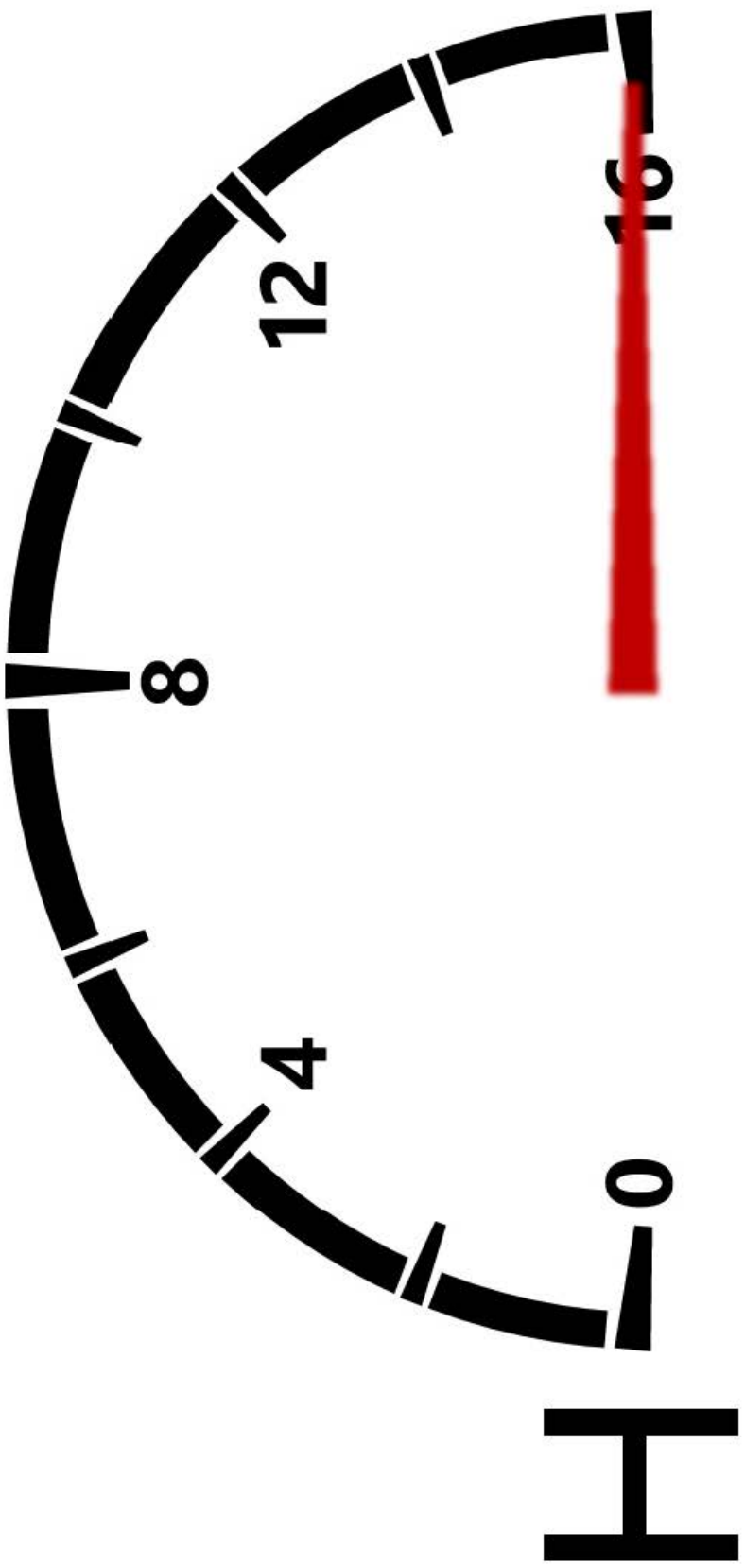
- Have you ever experienced anything like remotely controlling a robot before?
- Did you enjoy your experience?
- Do you see any of these interfaces being useful for remotely controlling robots?
- Can you think of other places or tasks where this interface would be useful?
- Do you think they are suitable for search and rescue?
- How would you change the interface to make the robot easier to control?
- What past experiences helped you the most when controlling the robot?
- What were the most memorable parts of your experience?
- Do you have any questions about the experiment or the technology?
- Do you have any last comments or questions?

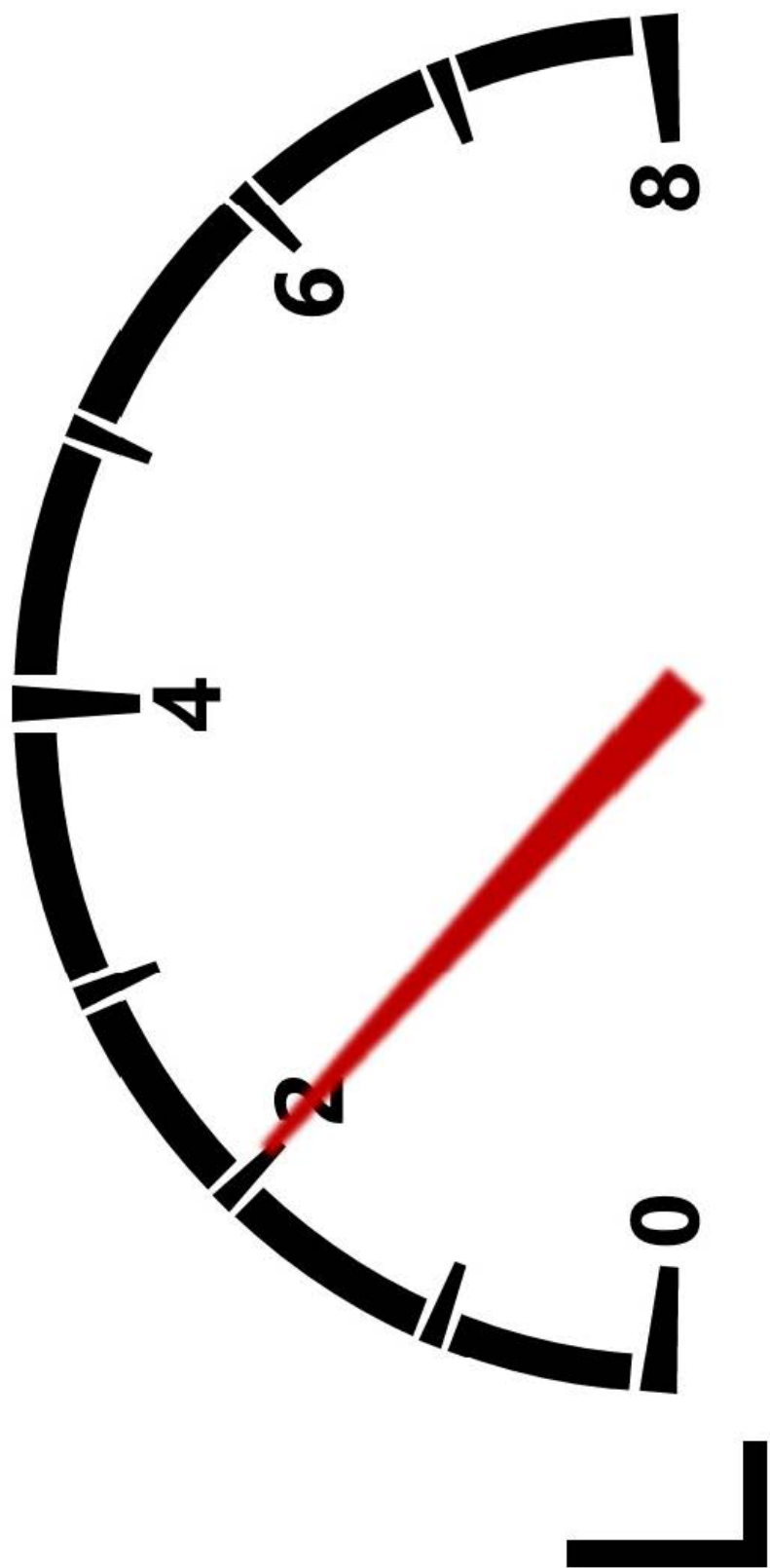
Disaster Area Meter Report

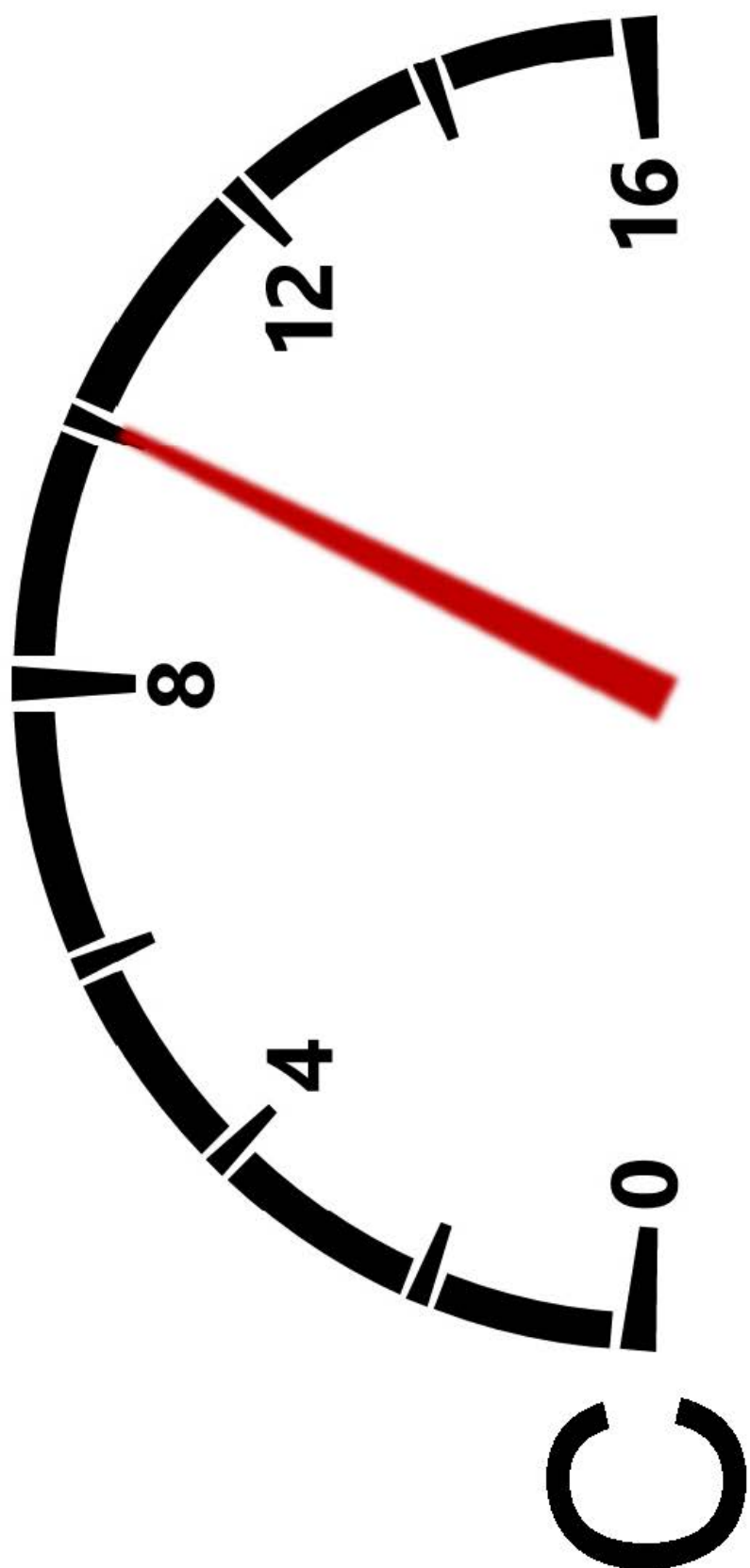
[illegible]

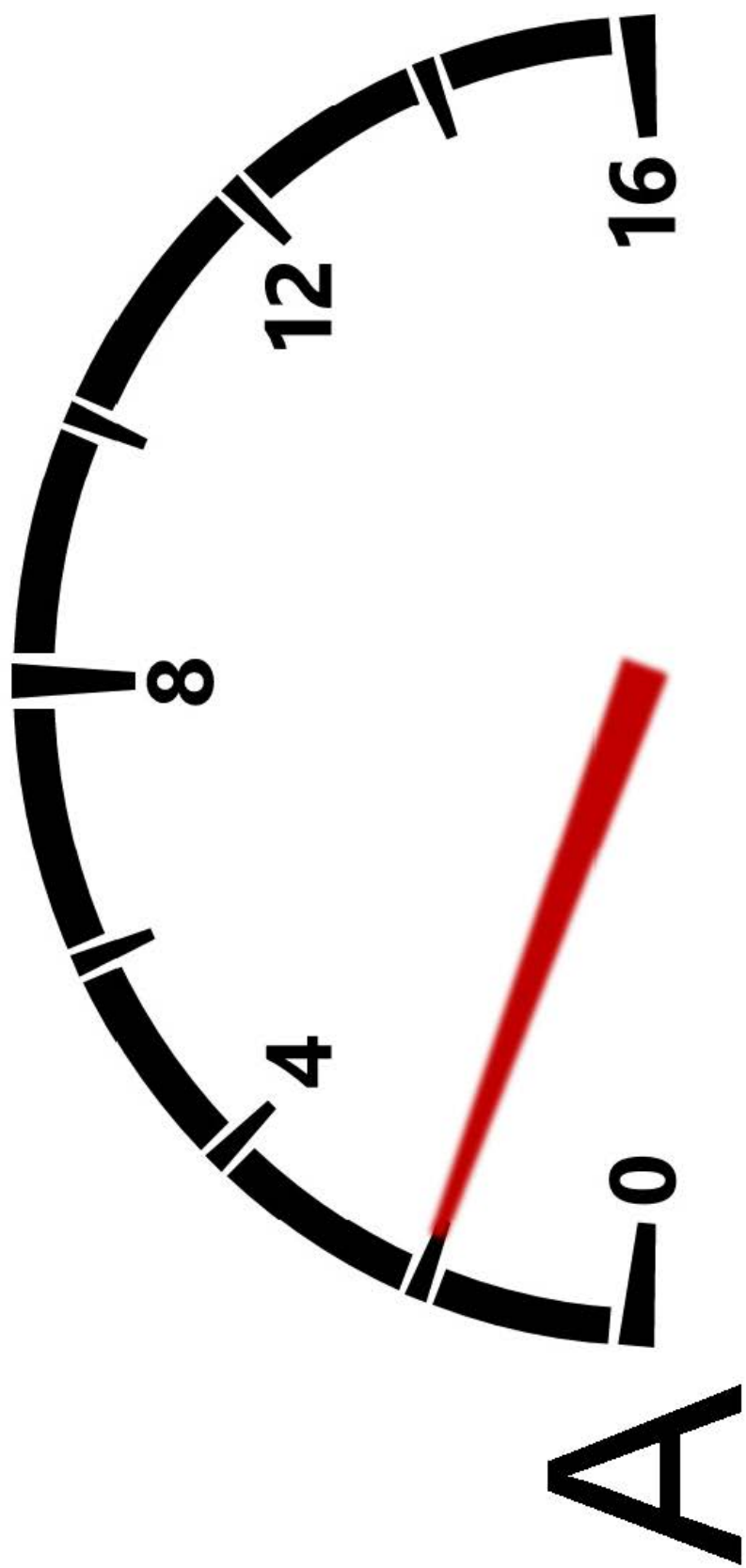


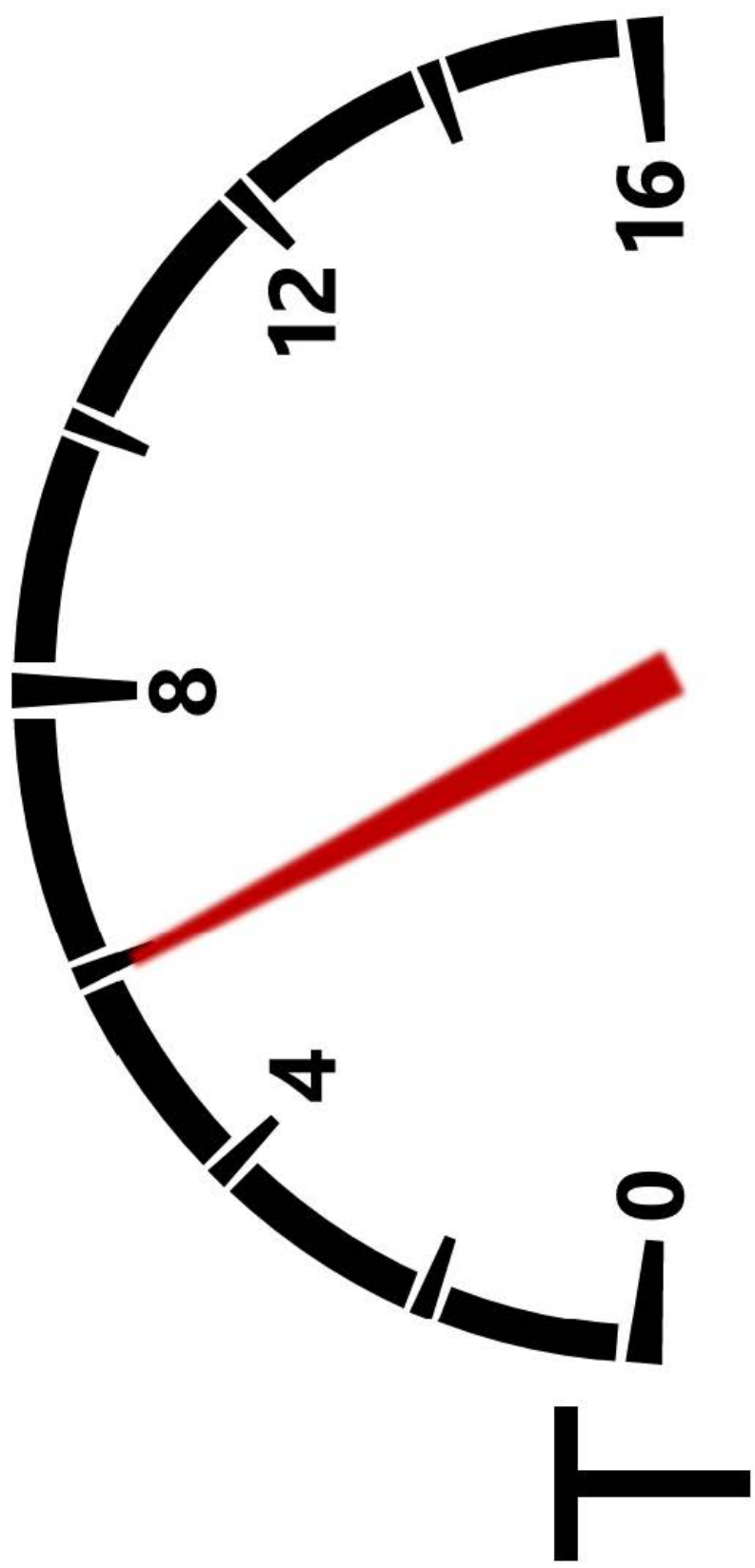


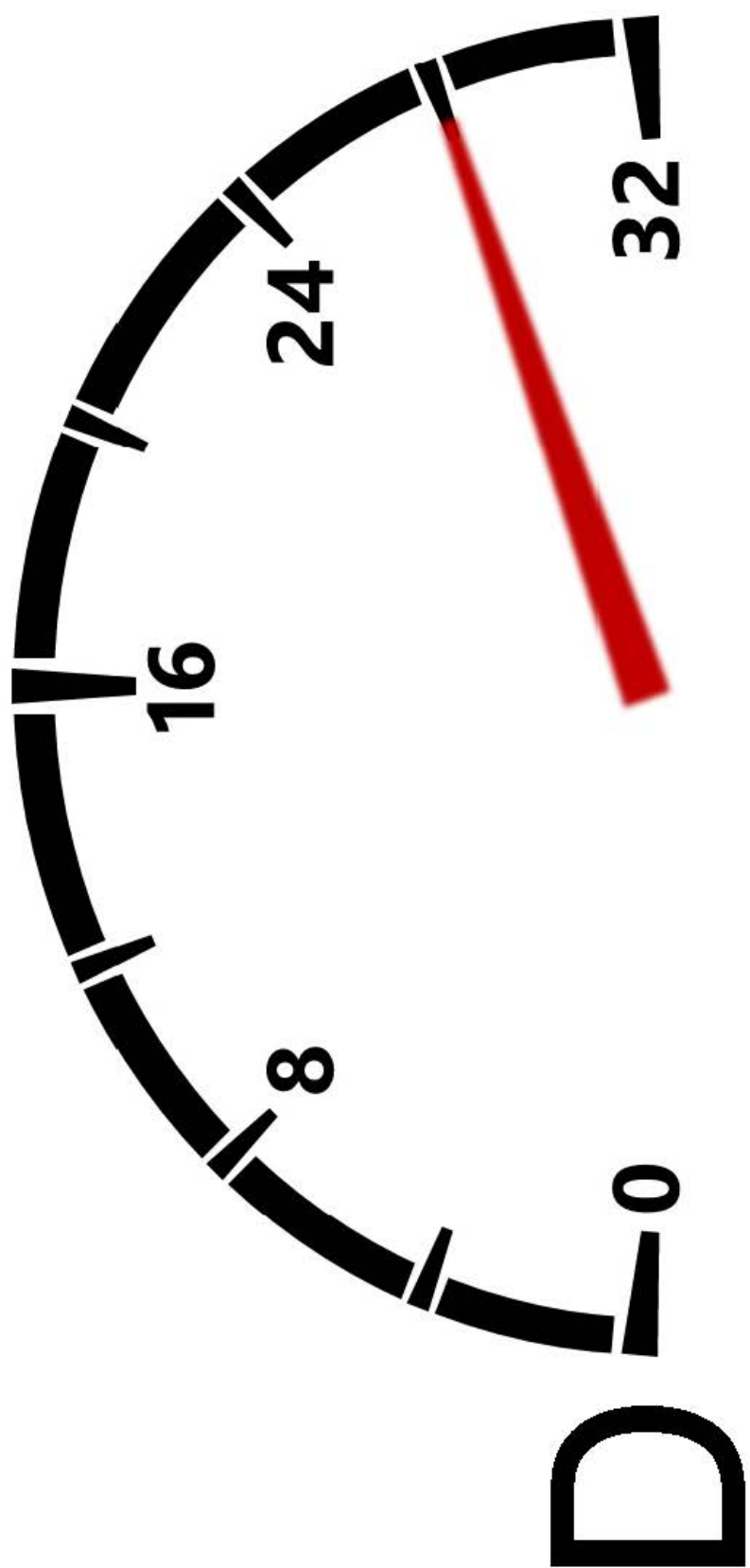


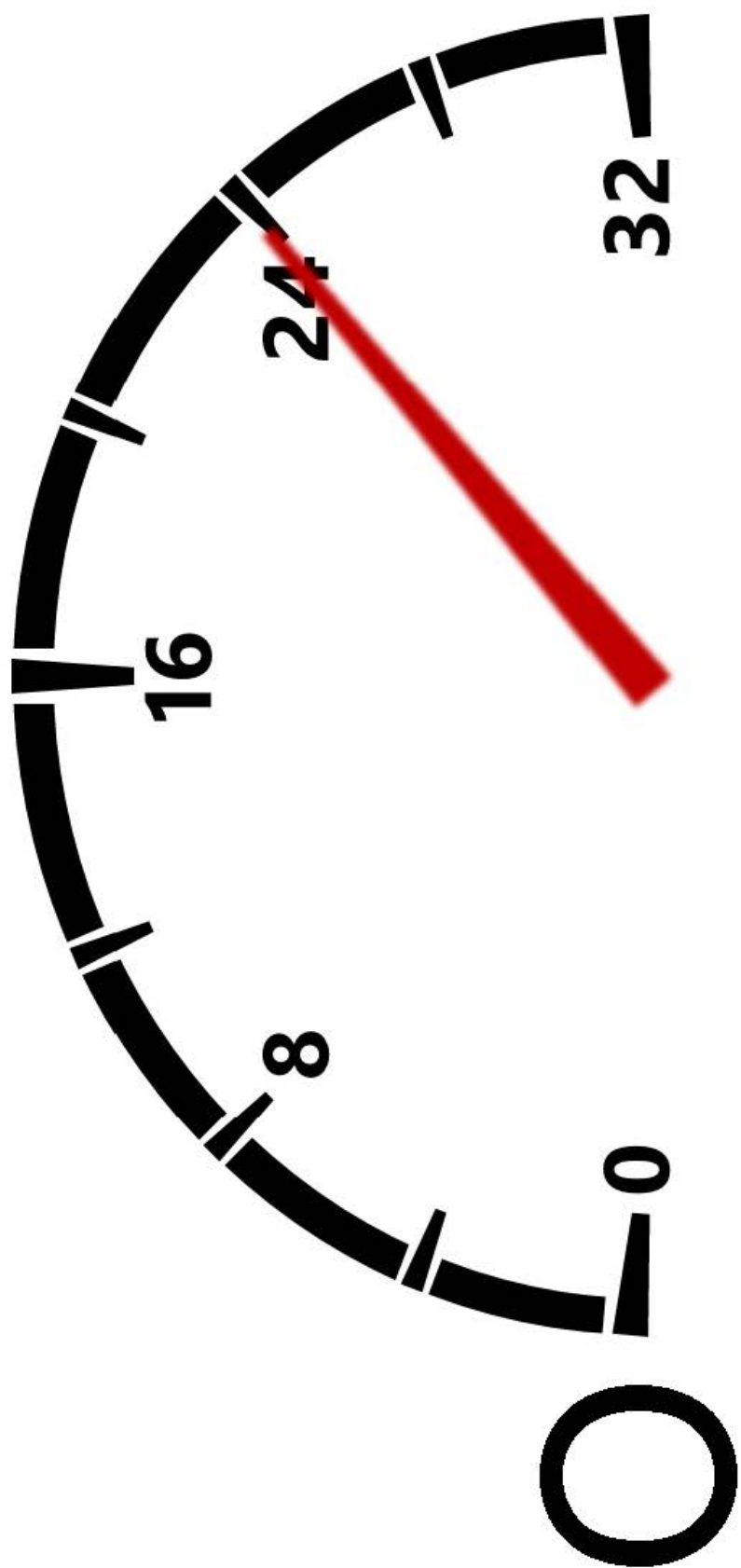


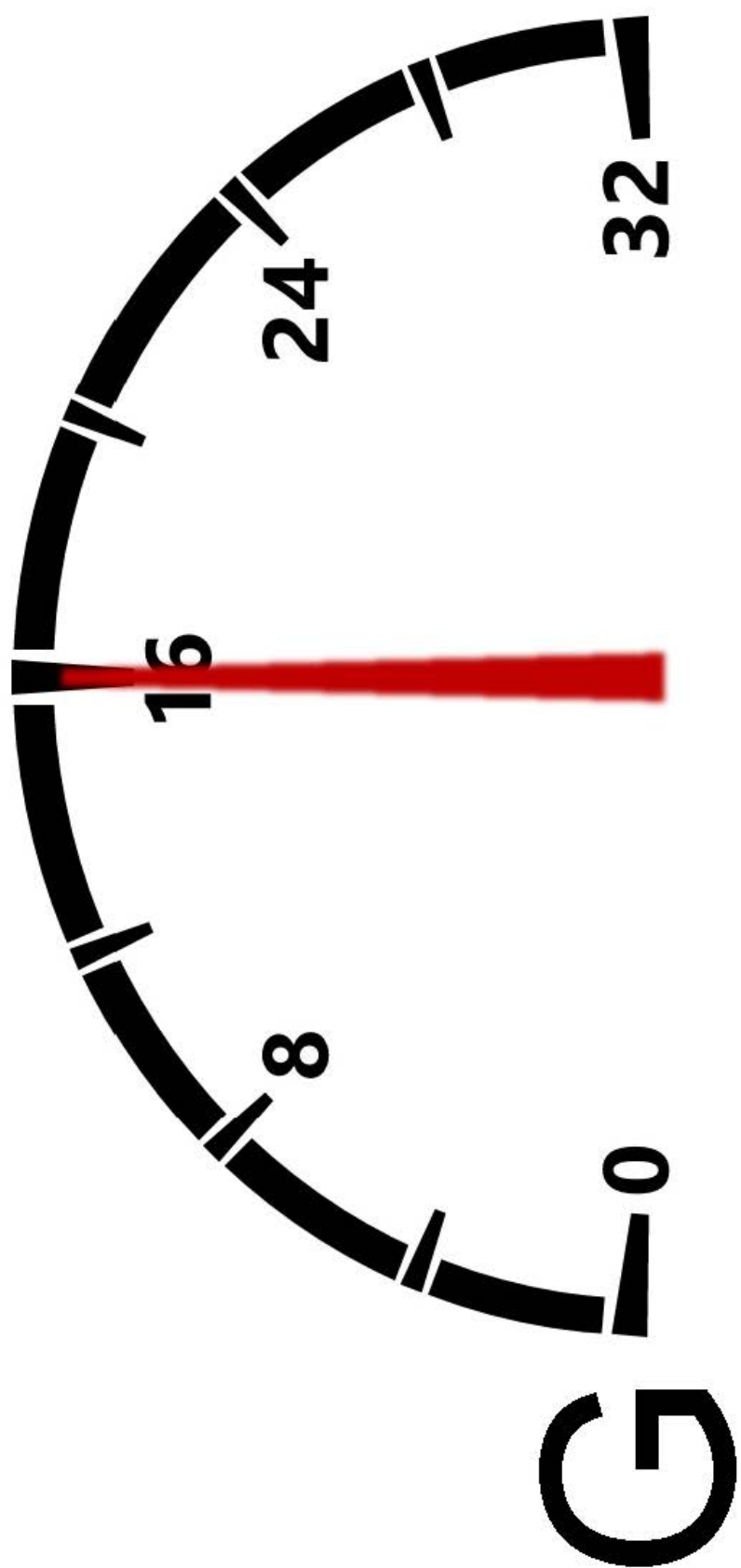










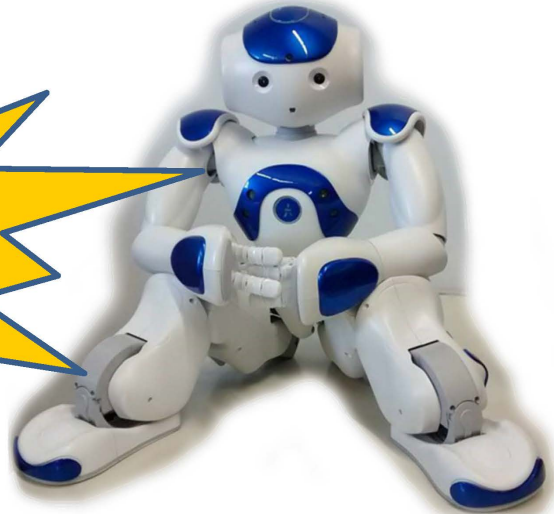


Appendix B Team Location: In-feed Embedded Techniques for Visualizing Robot Team Member Locations

- Public Recruitment Poster
- Consent Form
- Demographic Questionnaire
- Awareness Questionnaire
- Post-condition Questionnaire
- General Experience Questionnaire



UNIVERSITY
OF MANITOBA



Complete various tasks with NAO (a Humanoid Robot) in a one hour human-robot interaction experiment at the University of Manitoba. Note that you must be 18 or over to participate in our experiment.

Please visit:

[http://\[redacted\]](http://[redacted]) or

[http://\[redacted\]](http://[redacted])

If you have any questions about the study, please contact Stela H. Seo at [redacted] or Dr. James E. Young at [redacted].

This research study was approved by the Joint-Faculty Research Ethics Board, University of Manitoba. If you have any concerns or complaints about this project, you may contact the Human Ethics Secretariat at [redacted]

Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo



Stela H. Seo





Project Title: Improved interfaces for Robot Tele-operation

Researchers: Dr. James E. Young, Stela H. Seo, Daniel J. Rea

This consent form, a copy of which will be left with you for your records and reference, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

Participation in this study is voluntary, and will take approximately 60 minutes of your time. For this study, you will have the opportunity to interact with a humanoid robot (called NAO) to simulate aspects of search and rescue missions. NAO is a 58 cm tall humanoid (walking) robot. You will use camera feeds to remotely interact with the NAO robot. You will view the video from the robot via a computer. To begin, we will introduce you the NAO robot. We will provide instruction on how we expect you to interact with the robot. You will be given a number of tasks to assess the suitability of the interfaces for the NAO. No expertise or experience is necessary. You will receive \$15 for your participation.

All information you provide is considered completely confidential; your name will not be included, or in any other way associated, with the data collected in the study. Data collected during this study will be used for academic research and purpose of publication in an anonymous form. We may use anonymized video or audio data for purposes of public presentation and dissemination only with your express permission (given below). In addition, data will be retained for a maximum of five years in a locked office in the EITC building, University of Manitoba, to which only researchers associated with this study have access. Once published, results of the study will be made available to the public for free at <http://home.cs.umanitoba.ca/~young/>. Again, no personal information about your involvement will be included. Please note that the University of Manitoba may look at the research records to see that the research is being done in a safe and proper way.

Your signature on this form indicates that you have understood, to your satisfaction, the information regarding participation in the research project and agree to participate as a subject. By doing this you also confirm that you are of the age of majority in Canada (18 years or more). In no way does this waive your legal rights nor release the researchers, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from the study at any time, and to refrain from answering any questions asked, without prejudice or consequence. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

This research has been approved by the Joint-Faculty Research Ethics Board. If you have any concerns or complaints about this project, you may contact Dr. James Young at [REDACTED] or the Human Ethics Secretariat at [REDACTED]. A copy of this consent form has been given to you to keep for your records and reference.

For purposes of research and analysis it is necessary for the experiment to be videotaped.

Do you agree that any video footage taken may also be used for distribution of research, for example, through research videos or images taken from your video? If you say No, your video will be used for internal data analysis purposes only.

No ___ Yes ___ but only if you blur my face ___ AND/OR if you muffle my voice ___

Participant's Name _____ Signature _____ Date _____

Researcher's Name _____ Signature _____ Date _____

Demographics Questionnaire

1) What is your age?

2) What is your sex?

- ☐ Male
- ☐ Female
- ☐ Intersex

3) How often do you play 3D videogames such as shooters, racing games...?

- ☐ Never played videogames
- ☐ A few times a month or less
- ☐ Once a week
- ☐ More than once a week

4) How would you rate your current skill level for this kind of 3D video game?

- ☐ Very poor
- ☐ Poor
- ☐ Fair
- ☐ Good
- ☐ Very good

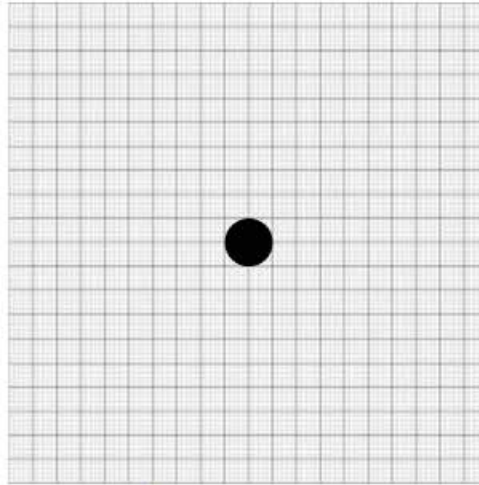
5) How would you rate your current vehicle driving skill level?

- ☐ Don't drive
- ☐ Very poor
- ☐ Poor
- ☐ Fair
- ☐ Good
- ☐ Very good

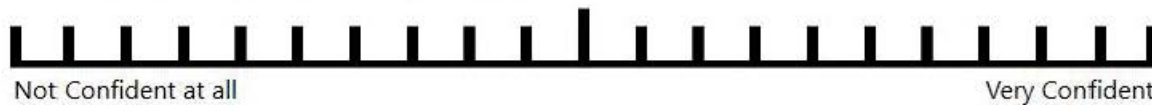
DONE

Awareness Questionnaire

1) Please indicate below where the other robots with respect to your robot?



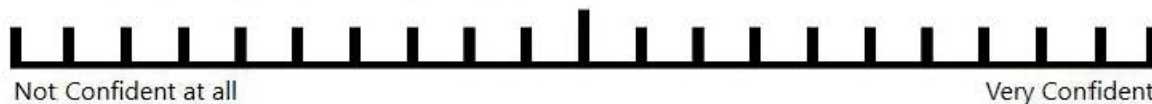
How confident are you on your response for the RED robot's location?



How confident are you on your response for the GREEN robot's location?



How confident are you on your response for the BLUE robot's location?



2) Which robot is the closest to your robot?

☐ RED ☐ GREEN ☐ BLUE

How confident are you in this answer?



2) Which robot is the farthest to your robot?

☐ RED ☐ GREEN ☐ BLUE

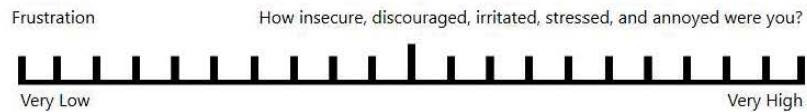
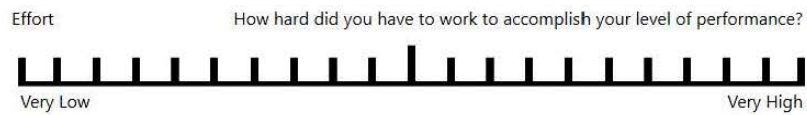
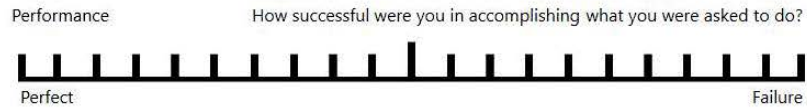
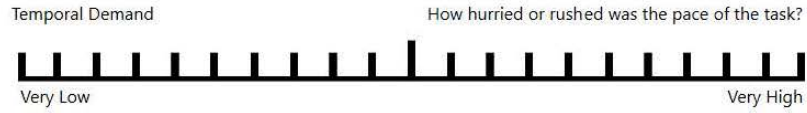
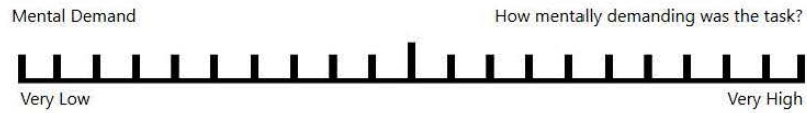
How confident are you in this answer?



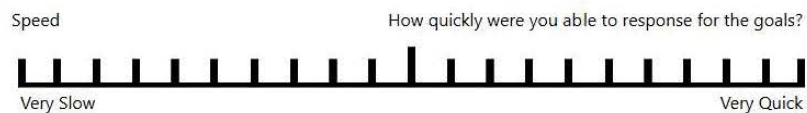
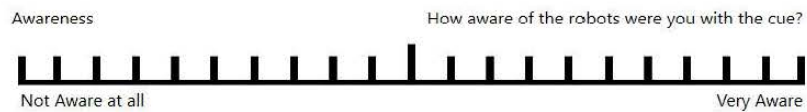
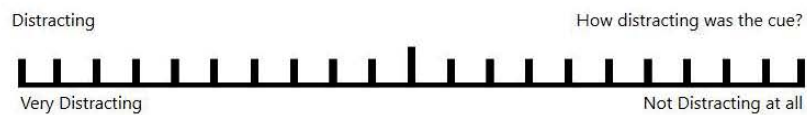
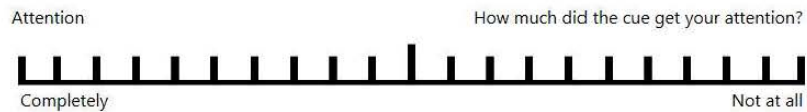
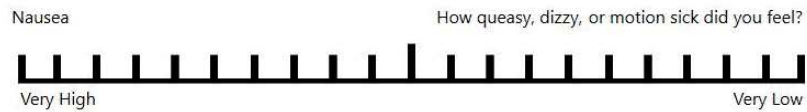
DONE

Post-condition Questionnaire

For the following question, mark ONE position along the scale:



For the following question, mark ONE position along the scale:



Please describe any pros and cons that you found for this cue:

Pros:

Cons:

DONE

General Experience Questionnaire

1) Please rank the interface cues from 1 (being the cue that you MOST preferred) to 4 (being the cue that

Map	<input type="text" value="0"/>
Halo	<input type="text" value="0"/>
Wedge	<input type="text" value="0"/>
Arrow	<input type="text" value="0"/>

2) Do you have any additional positive comments (if any)?

3) Do you have any additional negative comments (if any)?

4) Do you have any final comments or suggestions (if any)?

DONE

Appendix C Team States: Graphical Techniques for Maintaining Awareness of Team Robots

- Research Ethics and Compliance Approval Certificate
- Public Recruitment Poster
- Consent Form
- Demographic Questionnaire
- Awareness Questionnaire
- Post-condition Questionnaire
- General Experience Questionnaire

RENEWAL APPROVAL

Date: March 30, 2017

New Expiry: March 31, 2018

TO: James E. Young
Principal Investigator

FROM: Kevin Russell, Chair
Joint-Faculty Research Ethics Board (JFREB)



Re: Protocol #J2015:020 (HS17568)
“Exploring Alternate Camera Systems for Remotely Controlled Robots”

Joint-Faculty Research Ethics Board (JFREB) has reviewed and renewed the above research. JFREB is constituted and operates in accordance with the current *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*.

This approval is subject to the following conditions:

1. Any modification to the research must be submitted to JFREB for approval before implementation.
2. Any deviations to the research or adverse events must be submitted to JFREB as soon as possible.
3. This renewal is valid for one year only and a Renewal Request must be submitted and approved by the above expiry date.
4. A Study Closure form must be submitted to JFREB when the research is complete or terminated.

Funded Protocols:

- Please mail/e-mail a copy of this Renewal Approval, identifying the related UM Project Number, to the Research Grants Officer in ORS.



UNIVERSITY
OF MANITOBA

Earn \$15 and
contribute to
science!!



Complete various tasks with Jackal in a one and a half hours human-robot interaction experiment at the University of Manitoba. Note that you must be 18 years old or over to participate in our experiment.

Please visit:

[http://\[REDACTED\]](http://[REDACTED]) or

[http://\[REDACTED\]](http://[REDACTED])



If you have any questions about the study,
please contact Stela H. Seo at

[REDACTED]

or Dr. James E. Young at

[REDACTED].

This research study was approved by the Joint-Faculty Research Ethics Board, University of Manitoba. If you have any concerns or complaints about this project, you may contact the Human Ethics Secretariat at [REDACTED]



UNIVERSITY
OF MANITOBA

Project Title: Improved interfaces for Robot Teleoperation

Researchers: Dr. James E. Young, Stela H. Seo, Daniel J. Rea

This consent form, a copy of which will be left with you for your records and reference, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

Participation in this study is voluntary, and will take approximately 60 minutes of your time. For this study, you will have the opportunity to interact with a robot to simulate aspects of search and rescue missions. You will use camera feeds to remotely interact with the robot. You will view the video from the robot via a computer. To begin, we will introduce you to the robot. We will provide instruction on how we expect you to interact with the robot. You will be given a number of tasks to assess the suitability of interfaces for the robot. No expertise or experience is necessary. You will receive up to \$15 for your participation.

All information you provide is considered completely confidential; your name will not be included, or in any other way associated, with the data collected in the study. Data collected during this study will be used for academic research and purpose of publication in an anonymous form. We may use anonymized video or audio data for purposes of public presentation and dissemination only with your express permission (given below). In addition, data will be retained for a maximum of five years in a locked office in the EITC building, University of Manitoba, to which only researchers associated with this study have access. Once published, results of the study will be made available to the public for free at <http://home.cs.umanitoba.ca/~young/>. Again, no personal information about your involvement will be included. Please note that the University of Manitoba may look at the research records to see that the research is being done in a safe and proper way.

Your signature on this form indicates that you have understood, to your satisfaction, the information regarding participation in the research project and agree to participate as a subject. By doing this you also confirm that you are of the age of majority in Canada (18 years or more). In no way does this waive your legal rights nor release the researchers, sponsors, or involved institutions from their legal and professional responsibilities. **You are free to withdraw from the study at any time**, and to refrain from answering any questions asked, without prejudice or consequence. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

This research has been approved by the Joint-Faculty Research Ethics Board. If you have any concerns or complaints about this project you may contact Dr. James Young at [REDACTED] or the Human Ethics Secretariat at [REDACTED]. A copy of this consent form has been given to you to keep for your records and reference.

For purposes of research and analysis it is necessary for the experiment to be videotaped.

Do you agree that any video footage taken may also be used for distribution of research, for example, through research videos or images taken from your video? If you say No, your video will be used for internal data analysis purposes only.

No ___ Yes ___ but only if you blur my face ___ AND/OR if you muffle my voice ___

Participant's Name _____ Signature _____ Date _____

Researcher's Name _____ Signature _____ Date _____

Demographics Questionnaire

1) What is your age?

2) What is your sex?

☐ Male ☐ Female ☐ Intersex

3) How often do you play 3D video games in the first-person view (e.g., Counter-Strike, Overwatch, Skyrim, Need For Speed, etc.)?



4) How would you rate your current skill level for above 3D video games?



5) How often do you drive motor vehicle(s)?



6) How would you rate your current vehicle driving skill level?



7) Have you participated any of robot studies? If so, which research group's study and when did you participated, and what did you do in the study?

DONE

What was your team member doing?

Could you select the best description for your team robot?

The robot is:

☐ Moving

☐ Not Moving

☐ Looking Around

☐ Not Looking Around

The robot has:

Connectivity

☐ Strong Connectivity

☐ Okay Connectivity

☐ Weak Connectivity

Battery

☐ Strong Battery Level

☐ Okay Battery Level

☐ Weak Battery Level

Damage

☐ No Damage

☐ Light Damage

☐ Heavy Damage

Message

☐ No Message

☐ Message

☐ Urgent Message

How confident are you about the above answers?



Not at all

Very confident

Please describe the state of your team robot in your own words:

DONE

What was your team member doing?

Could you select the best description for your team robot?

The robot is:

☐ Moving

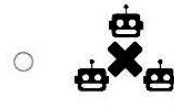
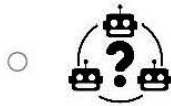
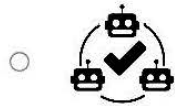
☐ Not Moving

☐ Looking Around

☐ Not Looking Around

The robot has:

Connectivity



Battery



Damage



Message



How confident are you about the above answers?



Not at all

Very confident

Please describe the state of your team robot in your own words:

DONE

What was your team member doing?

Could you select the best description for your team robot?

The robot is:

☐ Moving

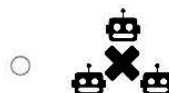
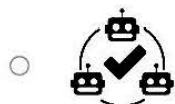
☐ Not Moving

☐ Looking Around

☐ Not Looking Around

The robot has:

Connectivity



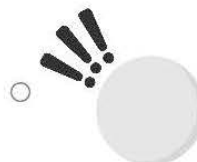
Battery



Damage



Message



How confident are you about the above answers?



Not at all

Very confident

Please describe the state of your team robot in your own words:

DONE

What was your team member on the left doing?

Could you select the best description for your team robot?

The robot is:

☐ Moving ☐ Not Moving

☐ Looking Around ☐ Not Looking Around

The robot has:

Connectivity
☐ Strong Connectivity ☐ Okay Connectivity ☐ Weak Connectivity

Battery
☐ Strong Battery Level ☐ Okay Battery Level ☐ Weak Battery Level

Damage
☐ No Damage ☐ Light Damage ☐ Heavy Damage

Message
☐ No Message ☐ Message ☐ Urgent Message

How confident are you about the above answers?



Please describe the state of your team robot in your own words:

What was your team member on the right doing?

Could you select the best description for your team robot?

The robot is:

☐ Moving ☐ Not Moving

☐ Looking Around ☐ Not Looking Around

The robot has:

Connectivity
☐ Strong Connectivity ☐ Okay Connectivity ☐ Weak Connectivity

Battery
☐ Strong Battery Level ☐ Okay Battery Level ☐ Weak Battery Level

Damage
☐ No Damage ☐ Light Damage ☐ Heavy Damage

Message
☐ No Message ☐ Message ☐ Urgent Message

How confident are you about the above answers?



Please describe the state of your team robot in your own words:

What was your team member on the left doing?

Could you select the best description for your team robot?



The robot is:

☐ Moving ☐ Not Moving

☐ Looking Around ☐ Not Looking Around

The robot has:

Connectivity

☐  ☐  ☐ 

Battery

☐  ☐  ☐ 

Damage

☐  ☐  ☐ 

Message

☐  ☐ 

How confident are you about the above answers?



Please describe the state of your team robot in your own words:

What was your team member on the right doing?

Could you select the best description for your team robot?

The robot is:

☐ Moving ☐ Not Moving

☐ Looking Around ☐ Not Looking Around

The robot has:

Connectivity

☐  ☐  ☐ 

Battery

☐  ☐  ☐ 

Damage

☐  ☐  ☐ 

Message

☐  ☐ 

How confident are you about the above answers?



Please describe the state of your team robot in your own words:

What was your team member on the left doing?

Could you select the best description for your team robot?

The robot is:

☐ Moving ☐ Not Moving

☐ Looking Around ☐ Not Looking Around

The robot has:

Connectivity

☐  ☐  ☐ 

Battery

☐  ☐  ☐ 

Damage

☐  ☐  ☐ 

Message

☐  ☐  ☐ 

How confident are you about the above answers?



Please describe the state of your team robot in your own words:

What was your team member on the right doing?

Could you select the best description for your team robot?

The robot is:

☐ Moving ☐ Not Moving

☐ Looking Around ☐ Not Looking Around

The robot has:

Connectivity

☐  ☐  ☐ 

Battery

☐  ☐  ☐ 

Damage

☐  ☐  ☐ 

Message

☐  ☐  ☐ 

How confident are you about the above answers?



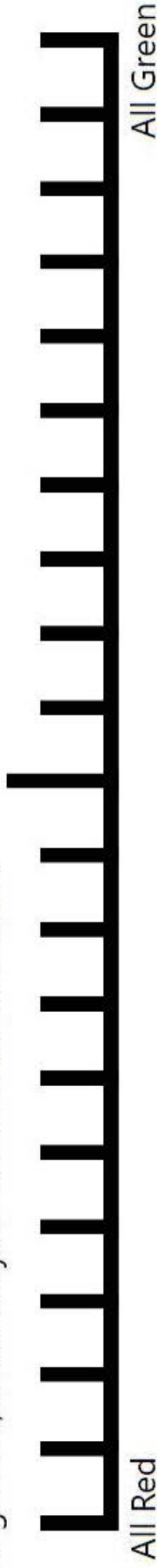
Please describe the state of your team robot in your own words:

Regarding your left team member. . .

How is your left teammate's connectivity?

- ☐ Strong Connectivity ☐ Okay Connectivity ☐ Weak Connectivity

In general, what are your left teammate's states?



Regarding your left team member. . .

How is your left teammate's battery level?

- ☐ Strong Battery Level ☐ Okay Battery Level ☐ Weak Battery Level

In general, what are your left teammate's states?

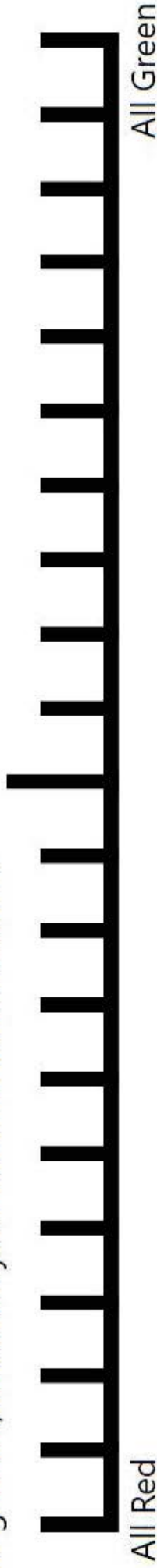


Regarding your left team member. . .

How damaged your left teammate?

- ☐ No Damage ☐ Light Damage ☐ Heavy Damage

In general, what are your left teammate's states?

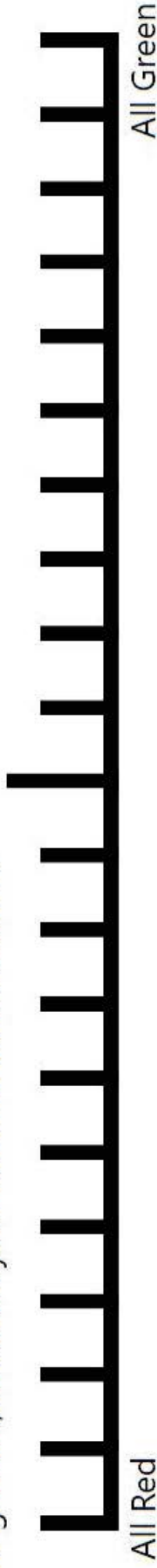


Regarding your left team member. . .

What is your left teammate's message?

- ☐ No Message ☐ Message ☐ Urgent Message

In general, what are your left teammate's states?

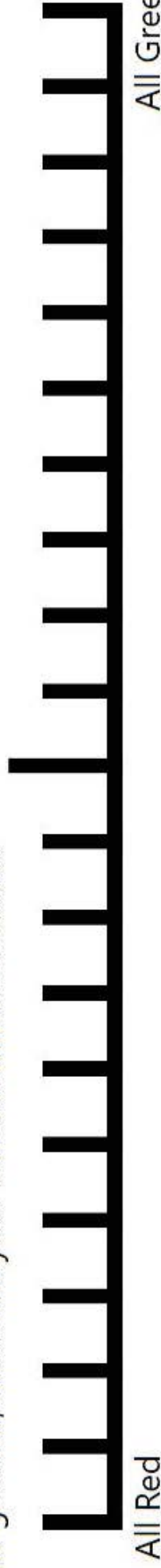


Regarding your left team member...

How is your left teammate's connectivity?

- ☐  ☐  ☐ 

In general, what are your left teammate's states?

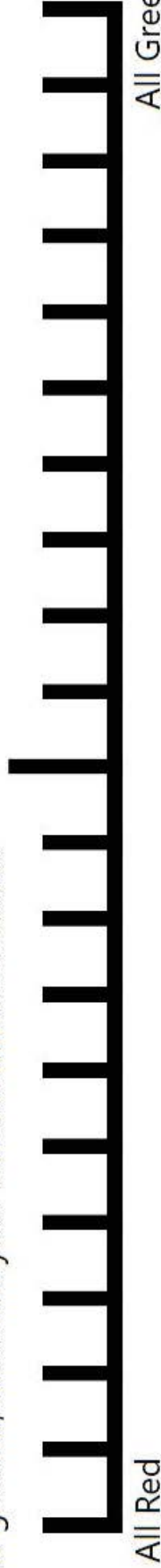


Regarding your left team member...

How is your left teammate's battery level?

☐ ☐ ☐ ☐

In general, what are your left teammate's states?

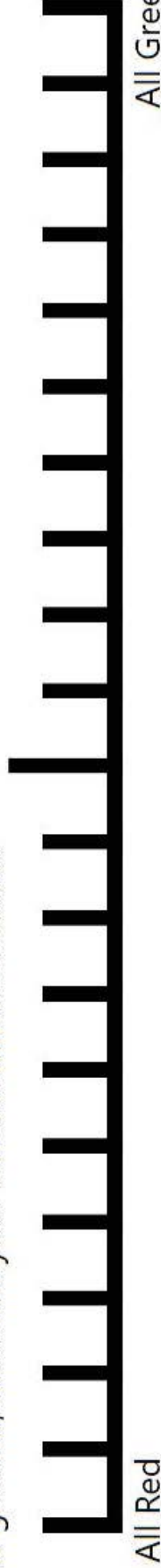


Regarding your left team member...

How damaged your left teammate?

- ☐ 
- ☐ 
- ☐ 

In general, what are your left teammate's states?

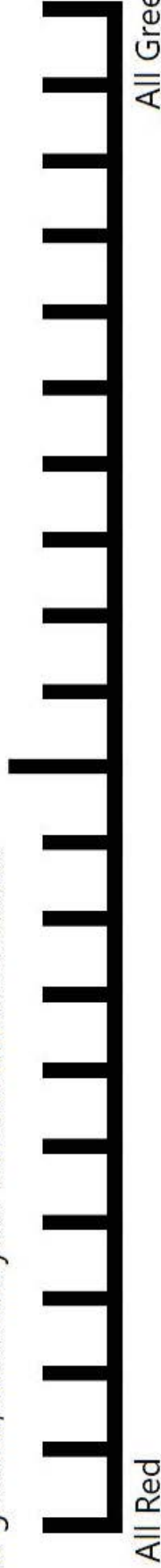


Regarding your left team member...

What is your left teammate's message?

- ☐ ?
- ☐ !

In general, what are your left teammate's states?

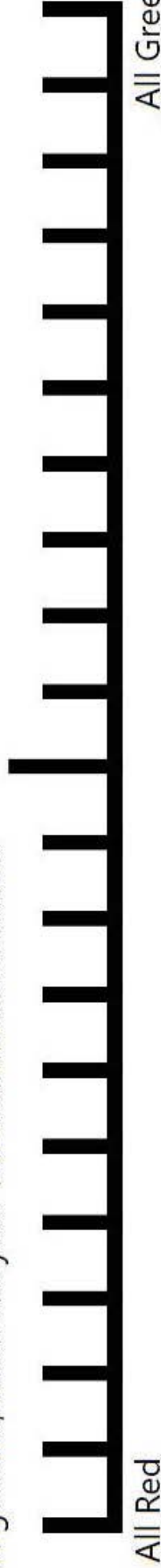


Regarding your left team member...

What is your left teammate's message?



In general, what are your left teammate's states?

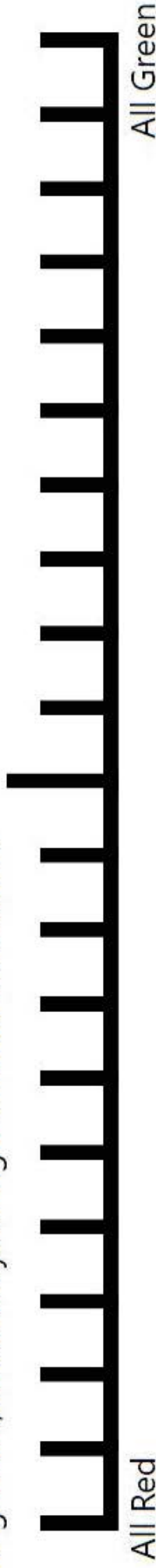


Regarding your right team member...

How is your right teammate's connectivity?

- ☐ Strong Connectivity ☐ Okay Connectivity ☐ Weak Connectivity

In general, what are your right teammate's states?



Regarding your right team member...

How is your right teammate's battery level?

☐ Strong Battery Level ☐ Okay Battery Level ☐ Weak Battery Level

In general, what are your right teammate's states?



Regarding your right team member...

How damaged your right teammate?

- ☐ No Damage ☐ Light Damage ☐ Heavy Damage

In general, what are your right teammate's states?

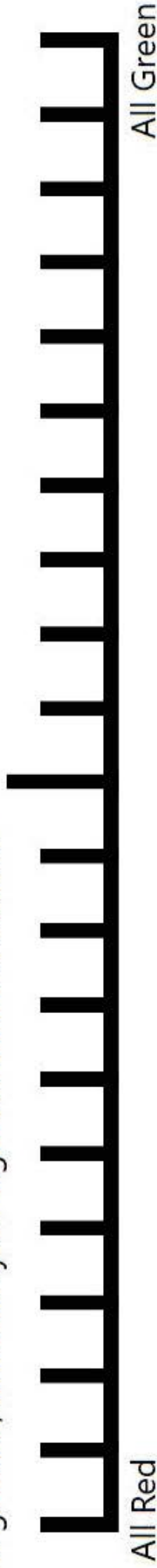


Regarding your right team member...

What is your right teammate's message?

- ☐ No Message ☐ Message ☐ Urgent Message

In general, what are your right teammate's states?

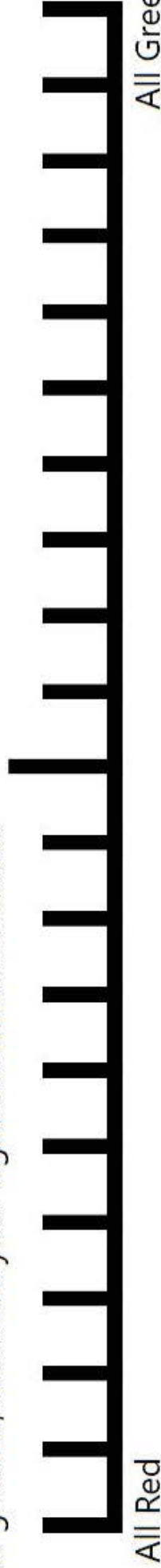


Regarding your right team member...

How is your right teammate's connectivity?

- ☐  ☐  ☐ 

In general, what are your right teammate's states?



Regarding your right team member...

How is your right teammate's battery level?

In general, what are your right teammate's states?



Regarding your right team member...

How damaged your right teammate?

- ☐  ☐  ☐ 

In general, what are your right teammate's states?



Regarding your right team member...

What is your right teammate's message?

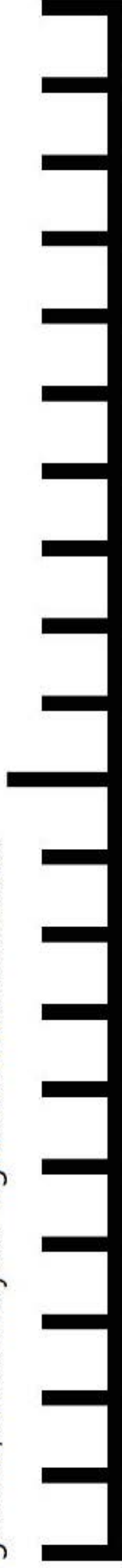
- ☐
- ☐

?

- ☐

!

In general, what are your right teammate's states?



All Red

All Green

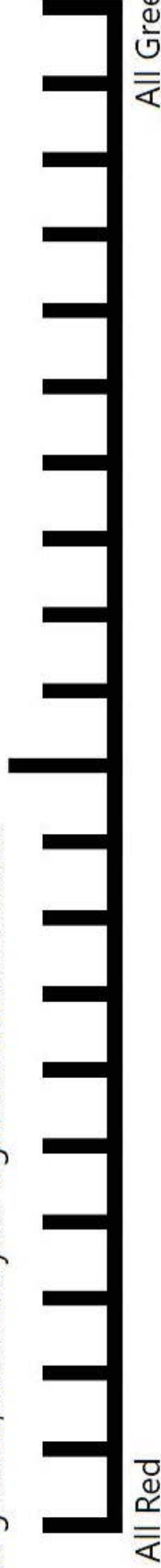
DONE

Regarding your right team member...

What is your right teammate's message?

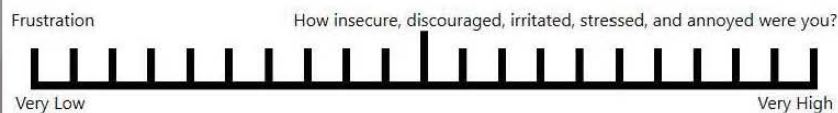
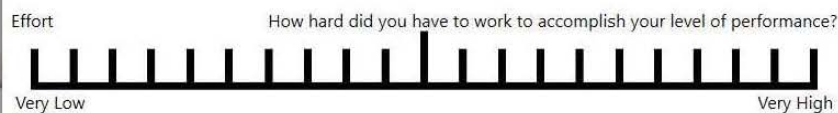
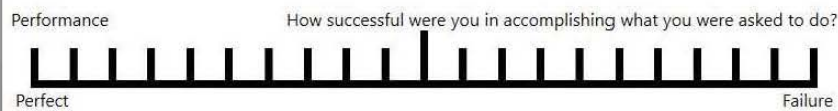
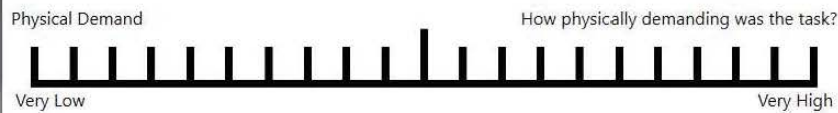
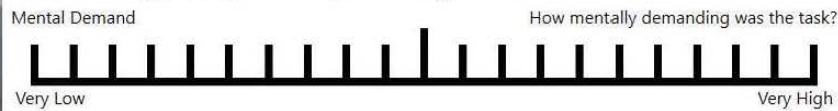


In general, what are your right teammate's states?

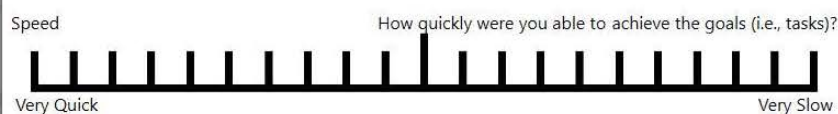
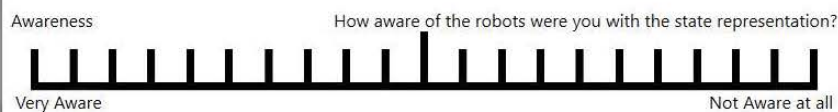
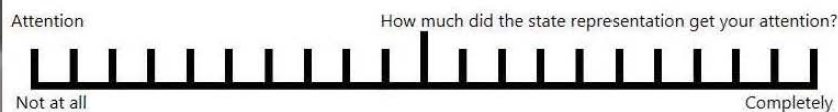


Post-condition Questionnaire

For the following question, mark ONE position along the scale:



For the following question, mark ONE position along the scale:



Please describe any pros and cons that you found for this representation (e.g., describe how easy or hard to understand the team robot's states using the representation, describe any idea to improve the representation):

Pros:

DONE

General Experience Questionnaire

1) Which state representation do you prefer?

☐ Text

☐ Simple Icons

☐ Icons with Emoji

2) Why do you prefer the representation over the others?

3) Did you find any similarities between representation techniques?

4) Did you find any differences between representation techniques?

5) Do you think your previous experience in gaming or driving helped you to accomplish the task?



Not at all

Definitely

And how so?

6) Do you have any additional positive comments (if any)?

7) Do you have any additional negative comments (if any)?

8) Do you have any final comments or suggestions (if any)?

DONE

Appendix D Control Transition: Utilization of Visual Transition as an Information Source in Multi-robot Teleoperation

- Research Ethics and Compliance Approval Certificate
- Public Recruitment Poster
- Consent Form
- Demographic Questionnaire
- Invitation Letter Draft



PROTOCOL APPROVAL

TO: James E. Young
Principal Investigator

FROM: Julia Witt, Chair
Joint-Faculty Research Ethics Board (JFREB)

Re: Protocol J2019:100 (HS23476)
“Design Workshops with Film and Media experts for Visual
Teleoperation Interface Designs”



Effective: January 9, 2020

Expiry: January 9, 2021

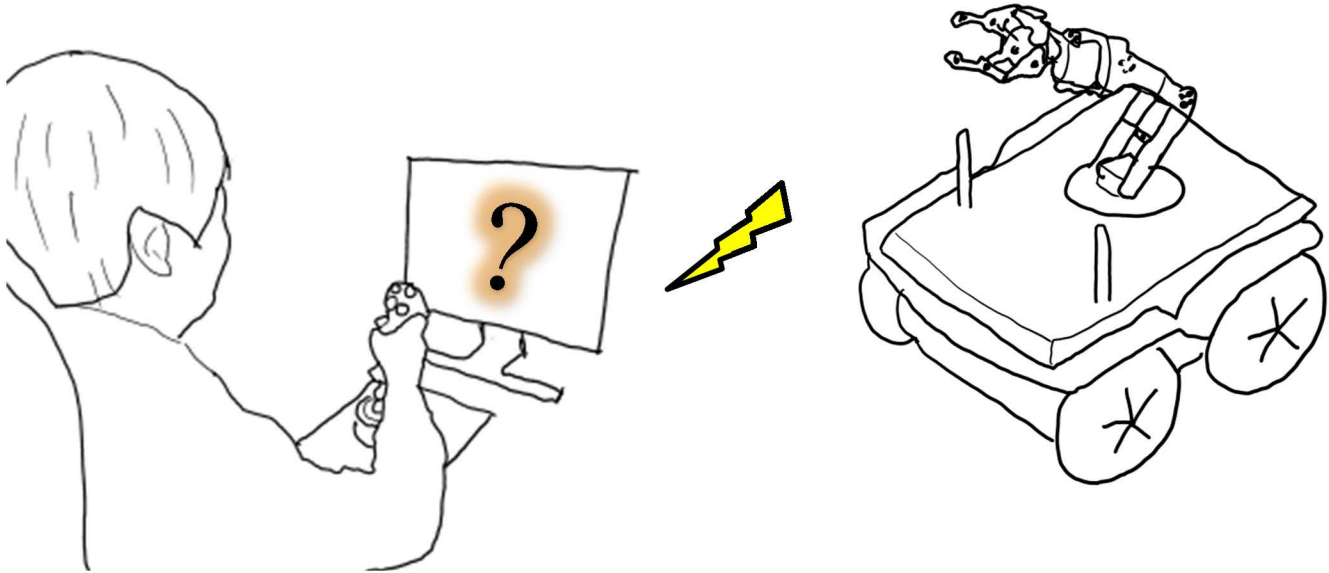
Joint-Faculty Research Ethics Board (JFREB) has reviewed and approved the above research. JFREB is constituted and operates in accordance with the current *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*.

This approval is subject to the following conditions:

1. Approval is granted for the research and purposes described in the application only.
2. Any modification to the research or research materials must be submitted to JFREB for approval before implementation.
3. Any deviations to the research or adverse events must be submitted to JFREB as soon as possible.
4. This approval is valid for one year only and a Renewal Request must be submitted and approved by the above expiry date.
5. A Study Closure form must be submitted to JFREB when the research is complete or terminated.
6. The University of Manitoba may request to review research documentation from this project to demonstrate compliance with this approved protocol and the University of Manitoba *Ethics of Research Involving Humans*.

Funded Protocols:

- Please e-mail a copy of this Approval, identifying the related UM Project Number, to the Research Grants Officer at [Redacted]



Interested in robots? Interested in user interfaces?

We invite creative people including film makers, media experts, artists, or students who are studying arts, cinematography, or video game designs to a design workshop for robot control interfaces.

Let's talk about it!

Come and chat in a 90-minute-long human-robot interaction workshop at the University of Manitoba and receive \$25 (and some snacks)! Note that you must be 18 years old or older to participate in this workshop, and the workshop will be video recorded for data analysis purposes.

To get details (e.g., date, time, and place) and participate in the workshop, please contact Stela H. Seo at [REDACTED].

This research study was approved by the Joint-Faculty Research Ethics Board, University of Manitoba. If you have any concerns or complaints about this project, you may contact the Human Ethics Secretariat at [REDACTED]

Project Title: Design Workshops with Film and Media experts for Visual Robot Interface Designs
Researchers: Dr. James E. Young, Stela H. Seo [REDACTED]

This consent form, a copy of which will be left with you for your records and reference, is the only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like to have more details, feel free to ask the onsite researcher. Please take the time to read this carefully and to understand any accompanying information.

Participation in this study is voluntary. Risks of participating in this study are no greater than in everyday life. The study is held in E2-551 EITC building in the University of Manitoba and is planned to take approximately 60-90 minutes of your time. For this study, you will be introduced to the problem of remotely controlling a robot or multiple robots. The researchers will introduce a design framework for creating new interface designs. You will help us by brainstorming and sketching new interface ideas. No expertise or experience is necessary. You will receive \$25 cash compensation for your participation at the beginning of the workshop. You can freely withdraw from the study at any time without any consequences (you can keep your honorarium) and choose not to answer any questions. However, please note that if you do withdraw, we are unable to delete your video data up until that point as it is a group setting. We will, however, destroy any sketches, prototypes, or notes that you have generated.

All information you provide is considered completely confidential; your name will not be included, or in any other way associated, with the data collected in the study. The video recording in this workshop is essential to capture the iterative design process (i.e., how people come up with an idea, create an initial design, and improve the design over time) and your body language used to explain your idea to others. If names used to address a certain person during the discussion, we will remove the part of the audio. Data collected during this study will be used for data analysis purposes. We may use anonymized quotes from the recording for purposes of public presentation and dissemination; however, we will not present video or screenshots. That is, your image will not be used in papers, presentations, put on the internet, etc. Once this workshop starts (i.e., video recording starts), there is no way we can remove the collected data. Please initial your response below. If you do not agree to be videotaped, please inform the researcher.

I DO consent to be videotaped for analysis purposes _____

Data will be stored in a locked office (E2-582) in the EITC building, the University of Manitoba, to which only researchers associated with this study have access until March 2025. Once published, results of the study will be made available to the public for free at hci.cs.umanitoba.ca. Again, no personal information about your involvement will be included. Please note that the University of Manitoba may look at the research records to see that the research is being done in a safe and proper way.

You have the option to sign up to receive a summary of findings for this research. This summary will be in non-scientific language and will be sent to you upon completion of our analysis by March 31, 2020. Please initial your response below if you want to receive the information:

I DO want to receive a summary of the findings of this research _____

If you do, please provide an email address or postal address

Your signature on this form indicates that you have understood, to your satisfaction, the information regarding participation in the research project and have agreed to participate as a subject. By signing the form, you also confirm that you are of the age of majority in Canada (18 years or more). In no way, this form waives your legal rights nor release the researchers, sponsors, or involved institutions from their legal and professional responsibilities. **You are free to withdraw from the study at any time**, and to refrain from answering any questions asked, without prejudice or consequence.

This research has been approved by the Joint-Faculty Research Ethics Board. If you have any concerns or complaints, you may contact Dr. James Young at [REDACTED], or the Human Ethics Secretariat at [REDACTED]. A copy of this consent form has been given to you to keep for your records and reference.

For purposes of research and analysis, it is necessary for the experiment to be videotaped. If you do not agree to be videotaped, please inform the researcher immediately.

Participant _____ Signature _____ Date _____

Researcher _____ Signature _____ Date _____

Demographics Questionnaire

1) What is your age?

2) What is your sex?

Male _____ Female _____ Other _____

3) What is your background relating to the arts? E.g., formal education, work experience, professional activities, performance history, etc.

4) How long have you studied/worked in this background?

Less than a year _____

1-3 years _____

3-5 years _____

5-10 years _____

10 years or more _____

5) Have you ever remotely controlled a robot or drone, etc?

Yes _____ No _____

If yes, what kind? _____

If yes, how would you rate your teleoperation skill level?


Not good at all Very good

Recruitment Email

Invitation to Visual Robot Interface Design Workshop

Hello, I am Stela H. Seo, a Ph.D. candidate in Computer Science at the University of Manitoba. I would like to invite creative people such as film makers, media experts, artists, or students who are studying arts, cinematography, or video game designs to participate in my visual robot interface design workshop. In this workshop, I will introduce a new framework for robot interface design, and as a group we will casually explore and sketch new ideas for remotely controlling robots. In short, I am exploring how to learn from standard techniques in the arts to develop better robot interfaces, and I need your help to explore my ideas.

The problem is that people are starting to remotely control a robot from a distant location, and sometimes, people even control multiple robots, or teams. However, it's quite challenging for the person to switch control between robots (and the rest of their work!), keeping all the information in their head (e.g., what was the robot doing? Where is it in the environment? What are my sensors telling me?). In my work, I've noticed that creative media tackles similar problems: creative writing, cinema, and theater, have developed techniques to help the audience understand where they are and all the relevant pieces of a story. I'm trying to see if we can learn from these techniques, for creating robot interfaces.

The workshop will be at most 90 minutes. You will be compensated for your time (\$25) and we will prepare refreshments (coffee, tea, and snacks). The workshop is planned to be held in E2-551 EITC, University of Manitoba. If you are interested in participating in this workshop, please contact me at [REDACTED] for scheduling and details.

For your information, the workshop will be fully video recorded for data analysis purposes; we will not present video or screenshots. That is, your image will not be used in papers, presentations, put on the internet, etc. You must also be the age of majority in Canada (18 years or more) to participate in this workshop. This research has been approved by the University of Manitoba Joint-Faculty Research Ethics Board; you may contact them if you have concerns or questions at [REDACTED].

Thank you for your time and consideration.

Please feel free to forward this recruitment to anyone you feel may be interested in our study.

Sincerely,

Stela H Seo, PhD Candidate, MSc
Human-Robot Interaction Lab
Computer Science
University of Manitoba