

ENHANCING A HUMAN-ROBOT INTERFACE  
USING A SENSORY EGOSPHERE

By

Carlotta A. Johnson

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

May, 2003

Nashville, Tennessee

Approved:

---

---

---

---

---

Date:

---

---

---

---

---

To God

be the glory for things he has done.

With him, all things are possible

and nothing is impossible.

Proverbs 3:5,6

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank God because this could not have been done without him. He placed the desire in my heart for a Ph.D. many, many years ago and I had no idea how this was going happen. Due to his grace and mercy, I have been able to accomplish what I thought was possible.

Second of all, I would like to thank my mother, Dorothy Johnson, for always being my greatest cheerleader. I thank her for always being there for me and giving me this desire to never stop learning. I also want to thank my natural family and my spiritual family for all of their support. These people include my natural brothers, Kevin and Kerry Johnson, and my spiritual sisters, Angela Lyles, Adija Rauls, Adrennia Hughley, Juvonda Hodge, Rosalind Robinson, Shawntelle Madison-Coker and Marlisa Johnson. I also want to thank my church family, Living Word Community Church for always keeping me in their prayers.

Finally, I thank Dr. Kazuhiko Kawamura, my dissertation adviser, for giving me the opportunity to work in the Vanderbilt University Intelligent Robotics Laboratory (IRL). This experience has been an awesome adventure and learning experience. I also want to thank Dr. Julie Adams for always making sure that I stayed on task and completed every task with a high standard of perfection and excellence. Additionally, I would like to thank the other members of my committee: Dr. R. Alan Peters, II, Dr. D. Mitch Wilkes and Dr. Nilanjan Sarkar. Also, I thank the many members of the IRL who made my tenure at Vanderbilt University a little less difficult, especially Tamara Rogers, Kim Hambuchen, Steve Northrup, Ed Brown, Jian Peng, Chai Nilas, Hande Keskinpala and Bugra Koku. Especially, Surachai Suksakulchai for always being so patient with me and teaching me how to program. I also thank Cordelia Brown for always giving me those words of encouragement at just those moments when I needed them the most. I also want to thank Dr. Shehu Farinwata, my first mentor and a primary motivator for me to return to school to pursue my doctorate. In closing, I thank you Lord for answering my prayers and giving me the desire of my heart.

## TABLE OF CONTENTS

|                                                           | Page |
|-----------------------------------------------------------|------|
| DEDICATION.....                                           | ii   |
| ACKNOWLEDGEMENTS.....                                     | iii  |
| LIST OF FIGURES .....                                     | viii |
| LIST OF TABLES.....                                       | viii |
| LIST OF ACRONYMS .....                                    | xiii |
| <br>Chapter                                               |      |
| I. INTRODUCTION.....                                      | 1    |
| Problem Statement.....                                    | 2    |
| Proposed Solution.....                                    | 3    |
| Outline .....                                             | 4    |
| II. HUMAN-MACHINE INTERFACES.....                         | 5    |
| Types of Interfaces .....                                 | 8    |
| Teleoperation Interfaces .....                            | 8    |
| Sensor Fusion Interface .....                             | 9    |
| Virtual Reality Interfaces.....                           | 15   |
| Novel Interfaces.....                                     | 18   |
| Types of Control .....                                    | 22   |
| Teleoperation.....                                        | 22   |
| Supervisory Control.....                                  | 24   |
| Collaborative Control.....                                | 27   |
| Applications .....                                        | 33   |
| III. EVALUATION OF HUMAN-MACHINE INTERFACES OVERVIEW..... | 35   |
| Measured Attributes.....                                  | 35   |
| Usability.....                                            | 36   |
| Mental Workload .....                                     | 38   |
| Situational Awareness.....                                | 41   |
| Evaluation Methods.....                                   | 43   |
| Heuristic Evaluation.....                                 | 43   |
| Cognitive Walkthrough .....                               | 44   |
| MUSiC .....                                               | 44   |
| SUMI .....                                                | 44   |

|     |                                                       |    |
|-----|-------------------------------------------------------|----|
|     | NASA-TLX .....                                        | 45 |
|     | Subjective Workload Assessment Technique (SWAT) ..... | 48 |
|     | Multiple Resource Questionnaire (MRQ) .....           | 48 |
|     | Spatial Reasoning .....                               | 49 |
|     | Contextual Inquiry .....                              | 49 |
|     | Cooperative Evaluation .....                          | 51 |
|     | Situational Awareness Rating Technique (SART).....    | 52 |
| IV. | SENSORY EGOSPHERE.....                                | 54 |
|     | Sensor EgoSphere .....                                | 55 |
|     | Sensory EgoSphere .....                               | 58 |
|     | Geodesic Dome Topology .....                          | 59 |
|     | Short-term Memory .....                               | 60 |
|     | Data Storage/Retrieval.....                           | 62 |
|     | Graphical Representation .....                        | 64 |
|     | Justification for the SES.....                        | 67 |
|     | Applications of the Sensory EgoSphere .....           | 67 |
|     | Supervisory Control.....                              | 67 |
|     | Egocentric Navigation .....                           | 71 |
| V.  | GRAPHICAL BASED HUMAN ROBOT INTERFACE.....            | 77 |
|     | Introduction .....                                    | 77 |
|     | Design Overview (Enhancing a HRI).....                | 83 |
|     | SES Agent .....                                       | 83 |
| VI. | EXPERIMENTAL DESIGN.....                              | 85 |
|     | Introduction .....                                    | 85 |
|     | Research Questions .....                              | 85 |
|     | Research Goal .....                                   | 86 |
|     | Research Hypotheses .....                             | 86 |
|     | Test Procedure.....                                   | 87 |
|     | Summary of Procedure .....                            | 87 |
|     | Place and Time.....                                   | 87 |
|     | Participants.....                                     | 87 |
|     | Equipment and Materials.....                          | 88 |
|     | Experimental Design .....                             | 88 |
|     | Data Collection.....                                  | 89 |
|     | Schedule.....                                         | 91 |
|     | Design Overview.....                                  | 92 |
|     | Situational Awareness.....                            | 94 |
|     | Task Scenarios .....                                  | 95 |
|     | Task One: Find the Robot (Training).....              | 95 |
|     | Task Two: Drive the Robot (Evaluation) .....          | 96 |

|       |                                                     |     |
|-------|-----------------------------------------------------|-----|
| VII.  | DATA ANALYSIS AND RESULTS .....                     | 99  |
|       | Participant Demographics .....                      | 99  |
|       | System Failures .....                               | 100 |
|       | Automatic Data Recording.....                       | 100 |
|       | Hardware.....                                       | 100 |
|       | Software.....                                       | 101 |
|       | Statistics .....                                    | 101 |
|       | Hypothesis Testing .....                            | 102 |
|       | Correlation .....                                   | 106 |
|       | Quantitative Results.....                           | 108 |
|       | Total Mouse Clicks.....                             | 108 |
|       | Task Scores .....                                   | 119 |
|       | Task Completion Times.....                          | 124 |
|       | Multiple Resources Questionnaire (MRQ).....         | 128 |
|       | NASA-TLX Workload Rating .....                      | 136 |
|       | Spatial Reasoning .....                             | 146 |
|       | Post-Task Questionnaire .....                       | 147 |
|       | Pre- versus Post-Experiment Display Comparison..... | 157 |
|       | Post-Experiment Questionnaire .....                 | 159 |
|       | Qualitative Results/ User Comments .....            | 162 |
|       | Training .....                                      | 162 |
|       | Task Reactions .....                                | 163 |
|       | Camera Display .....                                | 164 |
|       | Sonar, Laser, and Compass Displays .....            | 165 |
|       | Sensory EgoSphere Display .....                     | 166 |
|       | Landmark Map .....                                  | 167 |
|       | Robot Reactions.....                                | 168 |
|       | Interface Reactions .....                           | 169 |
| VIII. | DISCUSSION/CONCLUSIONS/FUTURE WORK.....             | 171 |
|       | Discussion .....                                    | 172 |
|       | Reduction in Task Execution Time .....              | 172 |
|       | Reduction in Mental Workload.....                   | 173 |
|       | Reduction in Participant Errors.....                | 175 |
|       | Increased Situation Awareness.....                  | 176 |
|       | Improved Usability .....                            | 177 |
|       | Conclusions.....                                    | 181 |
|       | Future Work.....                                    | 182 |
|       | Appendices                                          |     |
|       | A. NASA-TLX .....                                   | 185 |
|       | B. MULTIPLE RATINGS QUESTIONNAIRE (MRQ) .....       | 191 |
|       | D. PRE-EXPERIMENT QUESTIONNAIRE .....               | 203 |

|                                       |     |
|---------------------------------------|-----|
| E. POST-TASK QUESTIONNAIRE .....      | 206 |
| F. POST-EXPERIMENT QUESTIONNAIRE..... | 210 |
| G. TASK INSTRUCTIONS.....             | 218 |
| H. CONSENT FORMS .....                | 226 |
| I. EXPERIMENTAL RESULTS.....          | 235 |
| BIBLIOGRAPHY .....                    | 266 |

## LIST OF TABLES

| Table                                                                           | Page |
|---------------------------------------------------------------------------------|------|
| 2.1. Numeric relationships - Human-Robot ratios.....                            | 6    |
| 2.2. Spatial relationships - Intimacy and Viewpoint.....                        | 6    |
| 2.3. Authority Relationships [Murphy et al., 2001].....                         | 7    |
| 2.4. Levels of Mixed-Initiative Interaction .....                               | 28   |
| 3.1. Metrics for measuring usability .....                                      | 37   |
| 3.2. Relationship between SA and Workload [Endsley, 1995] .....                 | 41   |
| 3.3. Subscales of the NASA-TLX [Hart et al., 1998] .....                        | 46   |
| 3.4. RNASA-TLX rating scales [Cha et al., 1997] .....                           | 47   |
| 3.5. Ten-dimensional SART Scale [Selcon et al., 1989] .....                     | 53   |
| 4.1. Functional Classification of Memory Concepts [Kawamura et al., 2001b]..... | 73   |
| 5.1. Agent Structure of the ATRV-JR robot [Thongchai, 2001].....                | 78   |
| 6.1. Session 1 Schedule .....                                                   | 91   |
| 6.2. Session 2 Schedule .....                                                   | 91   |
| 6.3. Scenario One task allocation.....                                          | 96   |
| 6.4. Scenario 2 task allocation.....                                            | 98   |
| 7.1. Tasks A and B Camera Clicks (mean (m), standard deviations (s)).....       | 109  |
| 7.2. Tasks A versus Task B Camera Clicks.....                                   | 110  |
| 7.3. Tasks E and F Camera Clicks (mean (m), standard deviations (s)).....       | 111  |
| 7.4. Tasks E versus Task F Camera Clicks .....                                  | 112  |
| 7.5. Tasks A versus Task E Camera Clicks.....                                   | 112  |
| 7.6. Tasks B versus Task F Camera Clicks .....                                  | 113  |

|       |                                                                                  |     |
|-------|----------------------------------------------------------------------------------|-----|
| 7.7.  | Tasks E and F Map Clicks (mean (m), standard deviation (s)) .....                | 114 |
| 7.8.  | Tasks E versus Task F Map Clicks.....                                            | 115 |
| 7.9.  | Task B and Task F SES Clicks (mean (m), standard deviations (s)) .....           | 117 |
| 7.10. | Tasks B versus Task F SES Clicks .....                                           | 118 |
| 7.11. | Tasks A and B Scores (mean (m), standard deviations (s)).....                    | 120 |
| 7.12. | Tasks A versus Task B Scores.....                                                | 121 |
| 7.13. | Tasks E and F Scores (mean (m), standard deviations (s)).....                    | 122 |
| 7.14. | Tasks E versus Task F Scores.....                                                | 122 |
| 7.15. | Tasks A and B Completion Times (mean (m), standard deviations (s)).....          | 124 |
| 7.16. | Tasks A versus Task B Completion Times .....                                     | 125 |
| 7.17. | Tasks E and Task F Completion Times (mean (m), standard deviations (s)).....     | 126 |
| 7.18. | Tasks E and F Completion Times (Statistics).....                                 | 126 |
| 7.19. | Tasks A and Task B MRQ (mean (m), standard deviations (s)) .....                 | 129 |
| 7.20. | Tasks A versus Task B MRQ.....                                                   | 130 |
| 7.21. | Task E and Task F MRQ (mean (m), standard deviations (s)) .....                  | 131 |
| 7.22. | Task E versus Task F MRQ.....                                                    | 132 |
| 7.23. | Task Comparison MRQ.....                                                         | 133 |
| 7.24. | Tasks A and Task B NASA-TLX (mean (m), standard deviations (s)).....             | 137 |
| 7.25. | Tasks A versus Task B NASA-TLX .....                                             | 139 |
| 7.26. | Task E and Task F NASA-TLX (mean (m), standard deviations (s)).....              | 140 |
| 7.27. | Task E versus Task F NASA-TLX .....                                              | 141 |
| 7.28. | Task Comparison NASA-TLX.....                                                    | 142 |
| 7.29. | Task A and Task B Task Specific Ratings (mean (m), standard deviations (s))..... | 148 |
| 7.30. | Task A versus Task B Task Specific Ratings (Statistics) .....                    | 149 |

|       |                                                                                      |     |
|-------|--------------------------------------------------------------------------------------|-----|
| 7.31. | Task A and Task B General Questions (mean (m), standard deviations (s)).....         | 149 |
| 7.32. | Task A versus Task B General Questions (Statistics).....                             | 150 |
| 7.33. | Task A and Task B System Capability Ratings (mean (m), standard deviations (s))..... | 151 |
| 7.34. | Task A versus Task B System Capability Ratings (SPSS) .....                          | 152 |
| 7.35. | Task E and Task F Task Specific Ratings (mean (m), standard deviations (s)).....     | 152 |
| 7.36. | Task E versus Task F Task Specific Ratings (SPSS).....                               | 153 |
| 7.37. | Task E and Task F General Ratings (mean (m), standard deviations (s)) .....          | 154 |
| 7.38. | Task E versus Task F General Ratings (SPSS).....                                     | 154 |
| 7.39. | Task E and Task F System Capability Ratings (mean (m), standard deviations (s))..... | 155 |
| 7.40. | Task E versus Task F System Capability Ratings (SPSS).....                           | 156 |
| 7.41. | Task Comparison Statistics .....                                                     | 156 |
| 7.42. | Pre- versus Post-Experiment Comparison .....                                         | 158 |

## LIST OF FIGURES

| Figure                                                                                                 | Page |
|--------------------------------------------------------------------------------------------------------|------|
| 2.1. Sensor fusion user interface .....                                                                | 11   |
| 2.2. Sensor Fusion User Interface [Terrien et al., 2000] .....                                         | 13   |
| 2.3. Control Station with Virtual Environment .....                                                    | 18   |
| 2.4. Video (top left), command (top right), map (bottom left) and sensor<br>[Fong et al., 2001c] ..... | 19   |
| 2.5. Three-dimensional PC .....                                                                        | 21   |
| 2.6. Traditional teleoperation interface [Fong et al., 2001b] .....                                    | 23   |
| 2.7. Advanced teleoperation interface [Fong et al., 2001b] .....                                       | 24   |
| 2.8. Supervisory Control framework [Sheridan, 1992] .....                                              | 24   |
| 2.9. Teleoperation to Full Autonomy control Levels [Kortenkamp, 1997] .....                            | 31   |
| 2.10. Relationship between resources and task performance [Wickens et al., 1984] .....                 | 39   |
| 3.1. Usability Framework [Bevan, 1995a] .....                                                          | 38   |
| 3.2. Performance vs. Mental Workload [Burnett, 2001a] .....                                            | 40   |
| 3.3. Situation in the action-perception loop [Graefe, 1998] .....                                      | 42   |
| 3.4. Mental Rotation Sample Questions [Vandenberg et al., 1979] .....                                  | 49   |
| 4.1. Sensor EgoSphere for a Camera [Albus, 2001] .....                                                 | 57   |
| 4.2. Octahedron- and Icosahedron- Based Tessellated Dome .....                                         | 65   |
| 4.3. Relative position of Robot to SES [Johnson, 2002] .....                                           | 65   |
| 4.4. Sonar and Laser Ray Representation [Johnson, 2002] .....                                          | 66   |
| 4.5. Camera Options and Views for the Sensory EgoSphere [Johnson, 2002] .....                          | 67   |
| 4.6. Landmark EgoSphere .....                                                                          | 72   |
| 4.7. SES representation chain [Kawamura et al., 2002b] .....                                           | 74   |

|      |                                                                             |     |
|------|-----------------------------------------------------------------------------|-----|
| 5.1. | Sensor Suite on the ATRV-JR robot .....                                     | 81  |
| 5.2. | Multi-Agent-Based Robot Control Architecture [Kawamura et al., 2001a].....  | 82  |
| 5.3. | Integrated Agent-based Human-Robot Interface and Control Architecture ..... | 83  |
| 6.1. | Prototype of Human-Robot Interfaces .....                                   | 92  |
| 6.2. | Navigation Command Input Options .....                                      | 93  |
| 6.3. | Three Levels of Situation Awareness [Endsley, 1989] .....                   | 94  |
| 6.4. | Scenario One .....                                                          | 96  |
| 6.5. | Scenario Two.....                                                           | 97  |
| 7.1. | Display View Electronic Survey .....                                        | 157 |

## **LIST OF ACRONYMS**

|           |                                                               |
|-----------|---------------------------------------------------------------|
| AMAP      | Allocentric Map                                               |
| CI        | Contextual Inquiry                                            |
| DRUM      | Diagnostic Recorder for Usability Measurement                 |
| EMAP      | Egocentric Map                                                |
| GUI       | Graphical User Interface                                      |
| GUIM      | Graphical User Interface Manager                              |
| HCI       | Human-Computer Interface                                      |
| HRI       | Human-Robot Interface                                         |
| IMA       | Intelligent Machine Architecture                              |
| IRL       | Intelligent Robotics Laboratory                               |
| LES       | Landmark EgoSphere                                            |
| LTM       | Long-term Memory                                              |
| MCH       | Modified Cooper-Harper                                        |
| MIQ       | Machine Intelligence Quotient                                 |
| MMD       | Multi-Media-Display                                           |
| MMI       | Man-Machine Interface                                         |
| MRQ       | Multiple Resource Questionnaire                               |
| MUSiC     | Measurement of Usability in Context                           |
| NASA-TLX  | National Aeronautics and Space Administration Task Load Index |
| QUIS      | Questionnaire for User Interface Satisfaction                 |
| RNASA-TLX | Revised NASA-TLX                                              |
| SA        | Situation(al) Awareness                                       |
| SART      | Situational Awareness Rating Technique                        |
| SES       | Sensory EgoSphere                                             |
| SPM       | Sensory Processing Modules                                    |
| STM       | Short-term Memory                                             |
| SUMI      | Software Usability Measurement Inventory                      |
| SWAT      | Subjective Workload Assessment Technique                      |
| UIC       | User Interface Component                                      |
| WWL       | Weighted Workload                                             |

## **CHAPTER I**

### **INTRODUCTION**

This work presents a user study involving a human-robot interface using a discrete geodesic dome, called the Sensory EgoSphere. The Sensory EgoSphere is linked to the short-term memory database of a mobile robot. The memory database is searchable and is indexed by azimuth and elevation. This geodesic dome and its' associated database are called the Sensory EgoSphere (SES). It is proposed that the addition of the graphical geodesic dome portion of the SES to a human-robot interface will enhance usability as well as reduce the user's mental workload.

In supervisory control of mobile robots, it is sometimes difficult to determine the robot's present status when the supervisor is situated at a remote location. A remote supervisor is necessary in situations where there are environmental hazards or harsh working conditions. This chapter lays the foundation for a study to enhance a human-robot interface with the addition of a graphical representation of a robot's short-term memory structure, the Sensory EgoSphere (SES). This chapter introduces the problem statement and outline of this research on the SES and an enhanced Human-Robot Interface (HRI). This chapter describes the topics to be covered in the literature survey as well as the research outline. The chapter begins with a discussion regarding landmarks that may be essential for mobile robot autonomous or semi-autonomous navigation in known or unknown environments. The need to detect landmarks and use sensor readings to extract environmental information is a catalyst for the Sensory EgoSphere research.

Landmarks are distinct environmental features that a robot can recognize based upon sensory input. Landmarks can be geometric shapes with different colors or heights. In general, landmarks have fixed and known positions relative to which a robot can localize itself. Landmarks should be carefully chosen in order to be easily identified; for example, there must be sufficient contrast between the landmark and the background. A robot can navigate using landmarks, if the landmark characteristics are known and stored in the robot's memory. The primary localization task is to reliably recognize landmarks while calculating the robot's current position.

Landmarks are classified as either natural or artificial. Natural landmarks are those objects or features that exist in the environment and have a function other than robot navigation. Artificial landmarks are specially designed objects or markers that are specifically placed in the environment with the sole purpose of enabling robot navigation [Borenstein et al., 1996].

The Sensory EgoSphere is proposed as a viable solution to the coordination of distributed sensors in order to complete mobile robot navigation [Kawamura et al., 2002b]. The Sensory EgoSphere may also be used to enhance a human-robot interface by providing a robot-centric intuitive display of the robot sensory data [Johnson, 2002]. The addition of the SES to the HRI may facilitate supervisory control and increase user situational awareness while decreasing the user's mental workload.

### **Problem Statement**

In mobile robot applications it is often necessary for a human to teleoperate or supervise the mobile robot. An effective interaction technique involves the use of a graphical user interface (GUI). The user must be able to obtain a clear understanding of the

present robot status and environment in order to effectively supervise the mobile robot. Frequently the various sensory displays consume the interface display and the user is overwhelmed. The information disparity as well as the potential viewing modes complicates the user's ability to mentally consolidate the information.

### **Proposed Solution**

We hypothesize that, a graphical based HRI that incorporates the SES should provide a more intuitive sensory data display. This compact display is not considered to be sensor fusion, but rather a display that permits the user to mentally fuse notable events that occur in close proximity. Since the dome is centered on the robot frame, it provides the user with the robot's egocentric perspective or a semi-presence in the robot's environment. The SES graphic was designed using OpenGL ® in Visual Basic ® with the Intelligent Machine Architecture (IMA) [Pack, 1998]. The SES structure is a basic octahedron tessellated dome. Along with the graphical SES implementation, this research included the HRI and SES agent design. In order to evaluate the enhanced interface, a human factors study was performed. The study included 27 participants who performed four tasks over two days. The data collected included user satisfaction, perceived workload, task completion times and total mouse clicks. In order to teleoperate the robot via the graphical based interface, it was necessary to design basic robot behaviors. These basic behaviors included "move to point" and "move to object". The tasks were designed to have a scenario-based approach. Upon the completion of the usability study, the data was analyzed using a statistical software SPSS ®. The hypothesis was that the proposed enhanced HRI would decrease mental workload and increase situational awareness. Although there was a change in results based upon the interface participants used, these hypotheses were not upheld at a 5% level of significance.

This result implies a second phase of user studies with a larger population should be conducted.

### **Outline**

Chapter II provides a literature review of the current state of human-computer and human-robot systems. Chapter II also presents applications of man-machine interfaces in aviation, the military and space exploration. Chapter III presents methods for evaluating human-machine interfaces. The foundations and development of the Sensor and Sensory Egosphere (SES) are discussed in Chapter IV. Chapter IV also illuminates some of the applications of the Sensory EgoSphere. Chapter V highlights the design approach for an Intelligent Machine Architecture-based human-robot interface and the SES agent. Chapter VI provides the research methodology for the user evaluation and presents an explanation of the empirical study that was conducted. Chapter VII provides the quantitative and qualitative results of the human factors study. Finally, Chapter VIII contains the discussion, conclusions and future work.

## **CHAPTER II**

### **HUMAN-MACHINE INTERFACES**

This chapter presents a literature review covering the state of the art in human-computer and human-robot interfaces. Direct manipulation, sensor fusion, virtual reality and novel interfaces are discussed. Finally, some examples of interfaces in aviation, military and space exploration applications are presented.

A user interface provides the means by which humans and machines interact. Another term for user interface is a man-machine interface (MMI). The MMI includes all the components that the user encounters. The components include the input language, output language and interaction protocol. The term “human-computer interaction” was adopted in the mid-1980’s, and it describes a field of study that deals with all aspects of interaction between participants and computers [Preece et al., 1994]. In more recent years, HCI has been defined as "the discipline concerned with the design, evaluation and implementation of interacting computing systems for human use and with the study of major phenomena surrounding them" [Preece et al., 1994].

An experiment at Carnegie Mellon in 1997 involving the Dante II robot yielded the following guidelines for interface design [Bares, 97]. It must have:

1. Consistent appearance and interaction
2. Functional organization
3. Uncluttered layout
4. Simple command generation
5. Visual indication of safeguards

The primary goal of human-computer interaction is to design and implement effective and efficient user interfaces. Usability is often used as a measure of interface effectiveness. Since usability is not easily defined, the definition is usually developed through relation to cases and needs [Doherty, 2002].

Human-centered robotics emphasizes the study of humans as models for the robots or even the study of robots as models for humans. Three basic relationship taxonomies for human-centered robotics are numeric, spatial and authority. These three relationship taxonomies are given in Table 2.1, Table 2.2, and Table 2.3 [Murphy et al., 2001].

Table 2.1. Numeric relationships - Human-Robot ratios

| Humans      | Robots      |
|-------------|-------------|
| One person  | One robot   |
| One person  | Many robots |
| Many people | One robot   |
| Many people | Many robots |

Table 2.2. Spatial relationships - Intimacy and Viewpoint

| Role         | Human's Point of View | Spatial Relationship |
|--------------|-----------------------|----------------------|
| Commander    | God's eye             | Remote               |
| Peer         | Bystander             | Beside               |
| Teleoperator | Robot's eye           | Robo immersion       |
| Developer    | Homunculus            | Inside               |

Table 2.3. Authority Relationships [Murphy et al., 2001]

| Authority Relationship | Function      | Context Required              |
|------------------------|---------------|-------------------------------|
| Supervisor             | Commands what | Tactical situation            |
| Operator               | Commands how  | Detailed perception           |
| Peer                   | Cross-cueing  | Shared environment, functions |
| Bystander              | Interacts     | Shares environment            |

These authority relationships differ from the taxonomy defined by Scholtz [Scholtz, 2002] in the following ways. Scholtz also defines the supervisor, operator and peer relationships. These roles were expanded to include a mechanic and the peer role. These roles were split into a bystander and a teammate. The responsibility of the supervisor is to monitor and control the overall situation similar to Murphy's definition. The operator modified the internal software or models when the robot's behavior is not acceptable. The mechanic performs the physical changes to the robot in order to execute a desired behavior. The peer or teammate provides commands to the robot within the larger goal/intentions. The bystander is unable able to interact with the robot at the goal or intention level. The bystander only has access to a subset of the robot's full range of actions. Therefore, Scholtz's bystander and operator roles are the same as Murphy's. Scholtz has redefined the operator and peer as three roles that encapsulate all of the functionality of Murphy's two.

The three types of human-robot communication are direct, mediated and physical. Direct human-robot communication includes speech, vision, gesture and teleoperation. Mediated human-robot communication includes virtual environments and graphical user interfaces. The physical interaction includes mixed initiative and dialog based interaction [Murphy et al., 2001].

## **Types of Interfaces**

Human-robot interaction and especially human-robot communication is of primary importance for the development of robots that operate and cooperate with humans. This interaction is much more important when the robot operates outside of a manufacturing environment. Human-robot communication requires a user interface that allows the user to intuitively instruct the robot. This process involves translation of the user's intention into a correct and executable robot command. This process also requires useful and intuitive feedback so that the user immediately understands what the robot is doing. If the robot is to be operated by an inexperienced user, a higher level interface is necessary. Higher level interfaces must facilitate the human to robot communication that is tailored to the user and is easily understood. This chapter reviews interface considerations such as tele-manipulation, sensor fusion, virtual reality, and novel interfaces.

### **Teleoperation Interfaces**

In tele-manipulation interfaces, the operator directs the vehicle via hand-controllers while watching the video from vehicle mounted cameras. This type of interface is appropriate in the following situations [Fong et al., 2001a]:

1. Real-time human decision-making or control is required.
2. The environment can support high bandwidth and low-delay communication.

In the tele-manipulation of a remote vehicle, the vehicle is usually continuously displayed on the interface screen. Changes to the remote vehicle are represented and reversible. The user input typically involves a mouse, joystick, buttons, or touch screen.

Fong, when referring to telemanipulation, state, "Direct manipulation interfaces are easiest to apply to domains which permit concrete graphical representations" [Fong et al., 2001a].

### **Sensor Fusion Interface**

In many traditional teleoperation user interfaces, each part of the display is updated with data from a single sensor. Thus, the operator is forced to scan many display areas, interpret the information, and combine the results to obtain spatial awareness. For complex situations or a multi-sensor system, the resulting mental workload can be extremely high and leads directly to fatigue, stress, and inability to perform other tasks. Fusing the data from multiple sensors and presenting the result in a way that enables the operator to quickly perceive the relevant information for a specific task may reduce mental workload. This reduction of mental workload should free up the operator's mental resources to concentrate on the task itself.

The most difficult aspect of vehicle teleoperation is that the operator is unable to directly perceive the remote environment. The operator is forced to rely on sensors, bandwidth limited communications links, and an interface to provide all information. Based upon this difficulty, Fong [Fong et al., 2001c] has developed an approach that employs sensor fusion displays that combine information from multiple sensors or data sources to present a single, integrated view. This is necessary for applications in which the operator must rapidly interpret multispectral or dynamic data. In this research, an interface fuses sonar, lidar, and stereo range data.

Sensor fusion is commonly used to reduce uncertainty in localization, obstacle detection, and environment modeling. However, sensor fusion can also be used to improve teleoperation. Sensor fusion can create user interfaces that efficiently convey information,

facilitate understanding of remote environments, and improve situational awareness. Sensor fusion is accomplished by selecting complementary sensors, combining information appropriately, and designing effective representations.

Sensor fusion for teleoperation differs from classic sensor fusion because it considers human needs and capabilities [Meier et al., 1999]. In sensor fusion for robot teleoperation, the information the human may need, how it should be communicated, and how it will be interpreted must be identified. The appropriate method of combining information must be chosen. The same set of sensor data may be fused differently depending upon whether it is to be used by autonomous processes or by a human. For example, an environment modeling process may need multiple-sensor range data to be fused globally, but a human may only require local fusion of relevant sensor data. The representations must be effective so that the data is accessible and understandable. The interface display should simplify man-machine interaction. It is important to note that the addition of fused sensor data alone will not compensate for a poorly designed display.

Sonar is the common range sensor employed for sensor fusion. The advantage of sonar is that they can detect obstacles with high confidence. Since sonars provide active measurements, they are independent from the environmental energy. If an object is well defined, located perpendicular to the sonar axis, and has good ultrasonic reflectivity, a very precise range measurement can be obtained. One disadvantage of sonar ranging is that it is highly susceptible to error caused by non-perpendicular and off-axis targets. Range errors also occur due to multiple or specular reflections. In addition, sonar transducers have a wide beam cone that results in poor angular resolution. Meier [Meier et al., 1999] has developed a sensory fusion interface that takes advantage of the range sensing capabilities of the sonar sensor. Meier's interface contains two primary display areas,

1. a 2-D image with color overlay
2. a local map constructed with sonar data (see Figure 2.1)

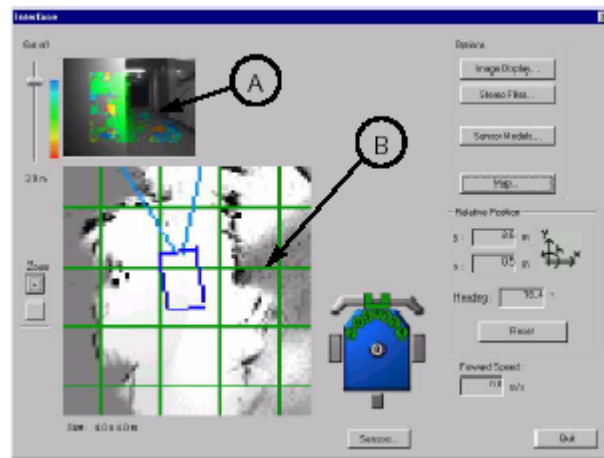


Figure 2.1. Sensor fusion user interface

The 2-D image is designed to facilitate scene interpretation and understanding. The color overlay directs the operator's attention to an obstacle located near the vehicle and aids distance estimation. The local map displays an occupancy grid, which is updated in real-time. It is hypothesized that the map designed will improve situational awareness and maneuvering in cluttered environments. The interface allows the operator to customize each display, i.e. color mapping, map scroll mode, display area, and display priority. The image display is created by overlaying range information as colors on a 2-D-intensity image taken from one of the cameras. It focuses the operator's attention on near objects, warns the operator if an object is very close, and enhances the estimation of relative distances [Terrien et al., 2000].

Stereo vision is employed primarily for range data since it has good angular resolution. The local map is built by combining vehicle odometry with stereo and sonar ranges on an occupancy grid. The advantage of this method is that it provides a very simple fusion process that updates a single, centralized map with each range sensor [Meier et al., 1999].

In Meier's work, sensor fusion was used to create displays that enable a better understanding of the remote environment while efficiently and accurately generating motion commands. This system was found to have weaknesses in certain environments due to the sonar/stereo combination. For example, smooth surfaces with low texture were frequently missed by both the sonars (specular reflection) and the stereo (poor correlation).

To address some of the system's sensing inadequacies, a SICK "Proximity Laser Scanner" (PLS) ladar was added to the sensor suite. Ladar sensors provide precise range measurement with very high angular resolution thus forming a good complement to the sonar and stereo sensors.

Terrien [Terrien et al., 2000] describes a remote-driving interface that contains sensor fusion displays and a variety of command generation tools. The interface is designed to improve situational awareness, facilitate depth judgment, support decision-making, and speed command generation. Considerable emphasis was placed on creating effective affordances and representations so that data is readily accessible and understandable. The operator generates remote driving commands by analyzing information displayed on the screen. Terrien intended to provide an interface that was intuitive, coherent, and maximizes information transfer. The main sensor fusion display contains three primary tools: the image display, the motion pad, and the map display. All the components of the sensor fusion display are shown in Figure 2.2.

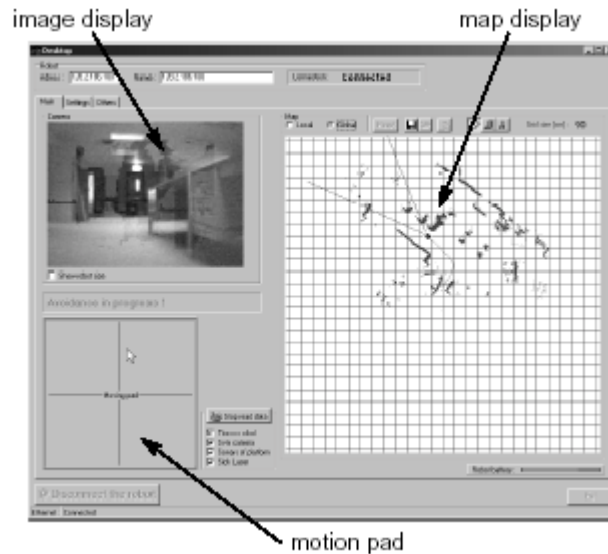


Figure 2.2. Sensor Fusion User Interface [Terrien et al., 2000]

The image display contains a monochrome video image with a color overlay to improve depth judgment and obstacle/hazard detection. The motion pad enables direct control of the robot. Clicking on the vertical axis commands a forward/reverse translation rate. Clicking on the horizontal axis commands a rotation rate. Translation and rotation are independent, thus the operator can simultaneously control both by clicking off-axis. The map display is employed to enable the supervisor to navigate the robot with a bird's eye view of the remote environment. The sensor fusion display is constructed as the robot moves and provides sensed environmental features as well as the robot's path. The map display provides both local and global maps. The local map provides the user the ability to precisely navigate through complex spaces. The global map allows large-area navigation while maintaining situational awareness. This map indicates where the robot has been. The operator can annotate the global map by adding comments or drawing "virtual" obstacles at any time.

Sensor fusion can be used to create displays that enable better understanding of the remote environment and to efficiently and accurately generate motion commands. The guiding principal in this work was that intelligent interfaces between humans and robots improve teleoperation performance. Truly integrated and efficient human-robot systems can only be achieved through the development of better interfaces [Terrien et al., 2000].

In 1990, Matsui and Tsukamoto developed a multi-media display for robot teleoperation [Matsui et al., 1990]. The multi-media display (MMD) allowed for the stereoscopic display of graphical models superimposed on real images taken from television cameras. The MMD is an advanced master-slave teleoperation system, where the robot performs autonomously and the operator makes only global decisions. The MMD includes superimposition, high-speed graphics, stereo-scope, and multiple windows. The three types of media that can be shown on the same screen are television images, three dimensional graphics, and text. Right and left images are shown on the screen and the operator wears stereoscopic glasses while viewing the screen. The graphics simulator is used to check for dangerous movements before the robotic manipulator executes a command. This system enables errors to be easily found by detecting the differences between superimposed images and the real environment and model [Matsui et al., 1990].

Another very similar approach to teleoperation was performed at the Jet Propulsion Laboratory in 1990. In this work, a predictive display was developed to teleoperate a 6-DOF PUMA robot. In this master-slave system, a phantom robot was controlled in real time and the image was updated on a high fidelity graphics display [Bejczy et al., 1990]. The image of the phantom robot on the monitor predicted the motion of the actual manipulator. The image of the robotic arm was overlaid on the actual camera view of the arm. Two-way communication allows force feedback to the user via the phantom robot. The simulated

image is overlaid with the delayed video image from the remote cameras that provides a real time simulated display of the manipulator and accurate displays of static objects. Preliminary experiments have shown that the predictive display enhances the human operator's telemanipulation task performance. The mean task completion time was reduced for all tasks completed [Bejczy et al., 1990].

### **Virtual Reality Interfaces**

Virtual Environments consist of highly interactive three-dimensional computer-generated graphics, typically presented to the user through a head-mounted, head-tracked stereo video display. Virtual environments completely replace reality by immersing the user inside a synthetic environment. Augmented reality allows the user to see the real world in combination with information superimposed or a composite on the display. Augmented reality enhances a user's perception of and experience with the real world.

### **Telepresence**

In an effort to achieve more efficient robot navigation, it may be necessary for humans to offer the best means of transmitting human problem solving and manipulative skills into hostile environments. Aiming at providing a "transparent" man-machine interface, some telepresence systems employ head mounted displays, sophisticated tracking sensors mounted on the operator's head and limbs, and force feedback. The ultimate goal of such systems is to make the human operator feel "present" at the remote site [Zhai et al., 1992].

Telepresence displays create an illusion of physical presence at the remote site. Telepresence is commonly claimed to be important for direct manual teleoperation, but the optimal degree of immersion required is still a topic for discussion [Meier et al., 1999].

The basic principle of telepresence is that if a robot can be interfaced with a human in a fashion that resembles human physical presence then high performance teleoperation can be achieved [Graves, 1998]. Experiential telepresence is the subjective feeling (mental state) of being at a remote place. Cybernetic telepresence involves the technological aspects of matching the characteristics of the robot element to those of the human operator [Graves, 1998]. One issue in this implementation is how to match remote robotic senses to the operator's human senses and human actions to specialized robotic actuators and manipulators. High fidelity telepresence systems require expensive and sophisticated input and output devices. In scenarios where both the environment is unstructured, there may be a requirement for the operator to perform some highly variable actions. High fidelity telepresence should lead to high performance teleoperation because participants feel as if they are present at the remote site. In this way, the participants can interact with the remote site as effortlessly and naturally as if they were actually there. Quantitatively, this should be characterized by a reduced workload level during teleoperation [Graves, 1998].

### **Augmented Reality**

Virtual reality techniques can be used to augment a human-computer interface. For example, computer graphics are used to provide alternative views that could not be achieved with live video. Lane [2002] presents a graphical simulation that was developed to allow an operator to visualize a telerobot and worksite in a three dimensional environment. Several windows are provided to simultaneously allow multiple views. Telemetry data, either from a training simulation or from sensors on the vehicles, are used to update the user interface and highlight changed states. Data from the simulations update the status on the control station panels. The graphical simulation is used in place of live data coming back from the robot.

Training simulations have been used to train new operators on the fundamentals of controlling the robot [Lane et al., 2002]. The operators learn how to properly use the different input devices, how each of the control station functions is utilized, and the procedure steps for specific tasks. Using these simulations, novices with no experience controlling robots, including young children, have learned enough to pilot buoyancy vehicles within a few minutes. The training simulation has quickly reduced the operator's learning time but the greater advantage of these simulations is that they have provided the capability to develop the robotic system. The graphical simulation replaces actual video during training. The graphical simulation can be used to augment live video during robotic operations. The capability to display the actual robot position within the graphical simulation has proven helpful in many circumstances. The ability to augment and even replace live video may improve operator's situational awareness. Furthermore, a graphical simulation has the advantage of displaying information that could never be observed from live video images. For large time delays, interactive real time control becomes difficult, and some level of supervisory or autonomous control becomes necessary. A predictive display can be used along with telemetry from the vehicle, thus allowing the user to see where the robotic system will be after the command motion is completed [Lane et al., 2002]. Figure 2.3 is an example of a control with a Virtual Environment.



Figure 2.3. Control Station with Virtual Environment

### **Novel Interfaces**

Novel interfaces employ non-typical input methods, output displays or are for unusual applications. This section reviews several of these interfaces, such as the PDADriver and the three-dimensional PC.

#### **PDADriver**

The first example of a novel interface is the PDADriver. The PDADriver enables remote driving anywhere and anytime using a Palm-size computer and low-bandwidth communication [Fong et al., 2000]. The problem with input devices such as joysticks or 2-D computer pointers is that the human-machine interaction is essentially static: the form and range of input is limited to physical devices. The most difficult aspect of remote driving, is that the operator is separated from the point of action. The driver must rely on information from sensors (mediated by communication links and displays) to perceive the remote environment. Consequently, the operator may fail to understand the remote environment and make judgement errors. This problem is most acute when precise motion is required. The PDADriver was designed to minimize the need for training, to enable rapid command

generation and to improve situational awareness. The four modes are video, map, command and sensors (see Figure 2.4 ).



Figure 2.4. Video (top left), command (top right), map (bottom left) and sensor [Fong et al., 2001c]

PDA's are attractive interfaces because they are lightweight, extremely portable and feature touch-sensitive displays. The drawback is that current PDA's have slow processors, limited memory/storage, and small displays. The PDA is used in command and control situations to direct the robot and to disambiguate natural language inputs [Fong et al., 2000].

With the PDA, remote driving is performed in a safeguarded, semi-autonomous manner, continuous operator attention is not required and the robot moves as fast as it deems safe. This should provide the operator with good situational awareness. Thus enabling the operator to switch between image and map displays, which was deemed invaluable. The operator also had the ability to understand what the robot was doing at a glance. Throughout this research, some improvements were suggested that would make it easier for the user to understand the remote environment, to better identify obstacles, and

areas to avoid. Some of these issues may be remedied with the addition of sensor fusion displays to the PDADriver. Fong's qualitative study showed that through a number of field tests, the interface had high usability and robustness. The participants stated that the interface caused them to have good situation awareness because they could rapidly switch between image and map displays. In this way, the user could understand at a glance what the robot was doing. The field tests provided only qualitative results.

### **Three-dimensional PC**

The three-dimensional PC is an example of a novel output display. Research efforts in computer science are concentrated on user interfaces that support the highly evolved human perception and interaction capabilities better than today's 2-D graphic user interfaces with a mouse and keyboard. Multimodal interaction not only makes working with a computer more "natural" and "intuitive" but also can substantially help to disambiguate the exchange of information in both directions between the user and the computer. The 3-D PC uses 3-D displays that do not require stereo glasses to present a 3-D graphic user interface. A newly developed 3-D display makes it possible to integrate the virtual interaction space into the real working space [Liu et al., 2000]. This system was tested with a general heuristic evaluation using usability experts. The three categories of evaluation were functionality, interface design and interaction. The participants rated the speech input particularly useful. The participants workload ratings were significantly lower for the gaze interaction.

Multimodal interface design attempts to incorporate a variety of human sensory and effector channels in combination. Humans are able to perceive information via multiple input channels using the senses of sight, hearing, touch, smell, taste, and balance. Perception via the visual and auditory modalities generally outperforms the haptic, olfactory and

gustatory modalities in terms of bandwidth, spatial and temporal resolution of the information transmitted. Figure 2.5 is a graphic of a person using the three-dimensional PC [Liu et al., 2002].



Figure 2.5. Three-dimensional PC

This work was based on the assumption that future participants of desktop computers will not tolerate encumbering interface devices attached to their head and body. The devices for the perceptual user interface as well as the 3-D display in this study rely on contact-free non-intrusive technologies. The mUltimo3-D system uses autostereoscopic 3-D displays creating the 3-D effect without the aid of polarizing glasses or any other headgear occluding the eyes [Liu et al., 2002].

### **Natural Language**

Skubic's [Skubic et al., 2002] work investigates the use of spatial relationships to establish a natural communication mechanism between people and robots. This research attempts to provide an intuitive interface that will be easy for novice participants to understand. In this work, a model of the environment was built by using sonar sensors on a mobile robot. From this model, a spatial description of the environment was generated. A

hand drawn map was then sketched on a PDA as a tool for communicating a navigation task to the robot. Spatial reasoning was used to analyze the map and it was combined with a natural language processing system. This combination provides the capability for the natural human-robot dialog using spatial language.

One proposed advantage of this interface is that the user can concentrate more on the task at hand rather than the modality of interaction. The user may employ various modes of input for commands and queries. The underlying goal of this work is to make robots capable of interacting freely with each other and with human participants.

This robot control system has been implemented as a distributed system with components for path planning, map processing, localization, navigation, and handling the various interface modalities, PDA, gesture, and input [Skubic et al., 2002].

### **Types of Control**

#### **Teleoperation**

Vehicle teleoperation is the act of operating a vehicle at a distance. It is used to operate vehicles in difficult to reach environments, to reduce mission cost and to avoid loss of life. Teleoperation can encompass any robot control from manual to supervisory control. Furthermore, the type of control may be shared/traded between operator and vehicle. Vehicle operation consists of three basic problems,

1. Where is the vehicle?
2. Where should the vehicle go?
3. How should the vehicle get there?

[Fong et al., 2001a]

Several characteristics distinguish vehicle teleoperation from remote control and other types of teleoperation. Teleoperation requires reliable navigation since vehicles are often deployed in unknown or unstructured environments. Also, vehicle teleoperation requires efficient motion command generation and calls for localized sensor data [Fong et al., 2001a]. Figure 2.6 is a traditional teleoperation interface with the human having sole responsibility for sensing and perception.

Figure 2.7 represents the advanced teleoperation interface. Note that instructions, control, and sensing can also take part on the machine side. This is possible because of collaborative control, which is addressed later in this chapter.

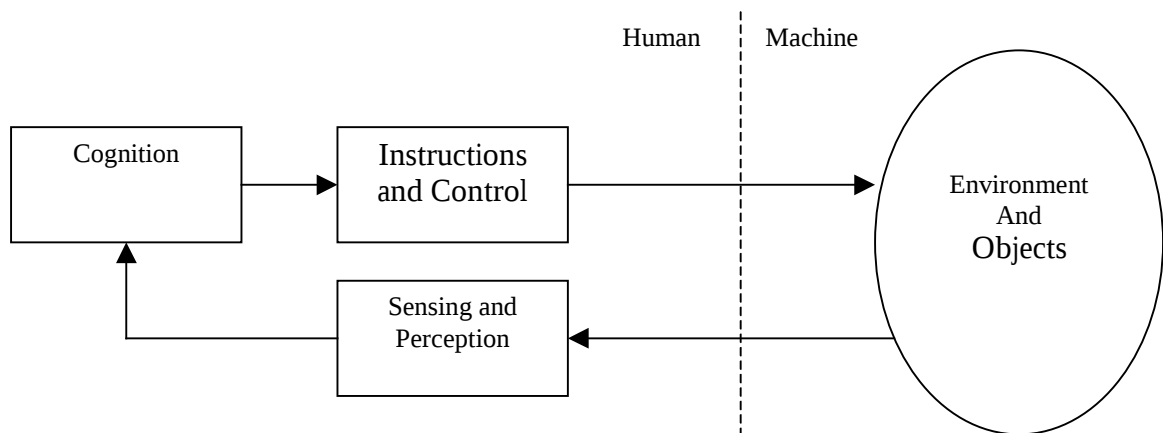


Figure 2.6. Traditional teleoperation interface [Fong et al., 2001b]

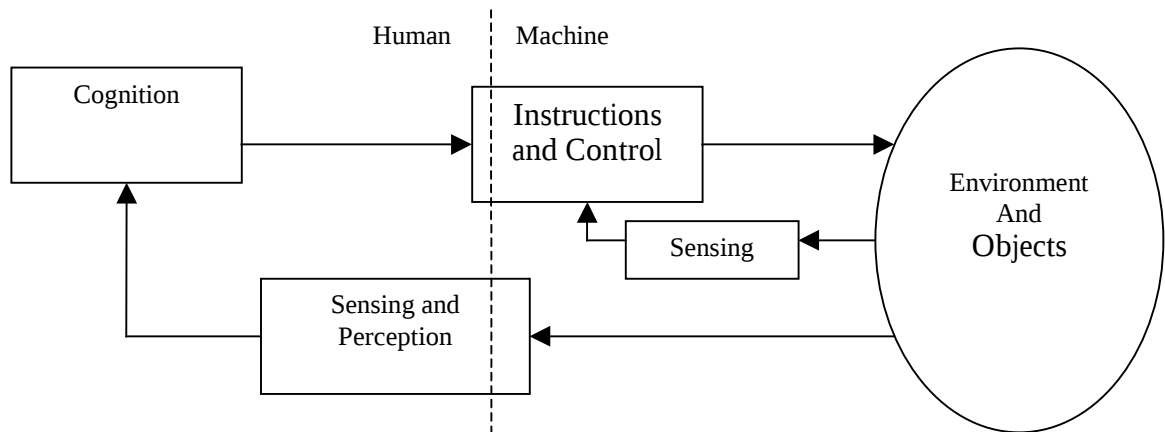


Figure 2.7. Advanced teleoperation interface [Fong et al., 2001b]

### Supervisory Control

Supervisory control requires a human-machine interface to permit the operator to monitor a machine and assist it if necessary. Under supervisory control, an operator divides a problem into a sequence of tasks, which the robot must achieve on its own. The steps to supervisory control are provided in Figure 2.8.

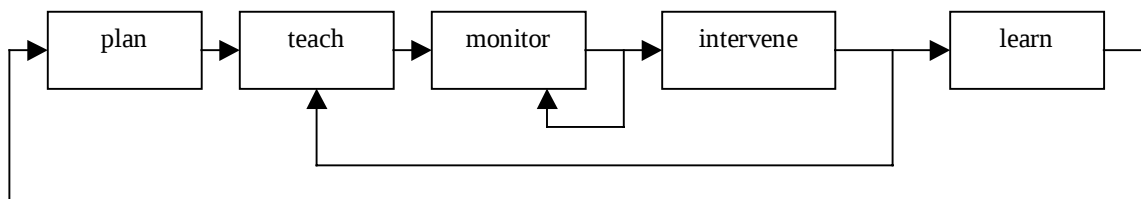


Figure 2.8. Supervisory Control framework [Sheridan, 1992]

In the realm of supervisory control, Murphy [1996] has designed a computer system that provides cooperative assistance for the supervision of remote semi-autonomous robots. This computer system consists of a blackboard-based framework that allows communication between the remote robot, the local human supervisor, and an intelligent mediating system. The intelligent mediating system aids interactive exception handling when the remote robot requires the assistance of the local operator.

A mobile robot must be able to perceive and move to perform tasks in environments where it is deemed too costly or dangerous for actual human presence. Since the technology has not yet produced a fully autonomous robot, there is still a strong need for human intervention. The interaction between human and robot is managed in a variety of ways collectively referred to as telesystems [Murphy, 1996]. Telesystems have the drawback of requiring high bandwidth communications in order for the human to perceive the environment and make corrections in the remote's action quickly enough [Murphy, 1996]. Even with adequate communication, the operator may experience cognitive fatigue due to the repetitive nature of many tasks, poor displays, and the demand of too much data and too many simultaneous activities to monitor. As robots use more sensors, the amount of data to be processed by the operator will increase, exacerbating the communication and fatigue problems and leading to less efficiency. The addition of artificial intelligence at the remote site is one solution to these shortcomings. "The intelligence involved in the operation of a mobile robot can be viewed as encompassing a continuous spectrum from master-slave teleoperation through full autonomy [Murphy, 1996]."

In Murphy's work, the intelligent sensing capabilities of a robot allow it to autonomously identify certain sensing failures and to adapt its sensing configuration. If the remote system cannot resolve the difficulty, it then requests assistance from the operator

through the teleVIA mechanism. This cooperative computerized assistant presents the relevant sensor data from other perceptual processes and a log of the remote robot's hypothesis analysis. This information is presented to the user in a form that can lead to an efficient and viable response. The local agent is composed of the human operator, together with a computational agent called the intelligent assistant that acts as an intermediary between the human and the robot. The intelligent assistant supports the perception and problem solving capabilities of the human and the robot by selectively filtering and enhancing perceptual data obtained from the robot. The intelligent assistant also assists in generating hypotheses about execution failures that cannot be solved by the remote robot. Each agent has internal routines called knowledge sources that read and post information to a global, asynchronous data structure called the blackboard. The operator, by definition, a knowledge source, communicates with the intelligent assistant and the remote robot via a graphical interface managed by the assistant. The display may contain different types of images obtained from various sensors involved in the failure, as well as textual information on the hypotheses generated and tested through the robot's autonomous exception handling mechanism. The development of the cooperative system has a number of specific goals [Murphy, 1996]:

1. “improve the speed and quality of the system's problem solving performance
2. reduce cognitive fatigue by managing the presentation of information
3. maintain low communication bandwidths by requesting only relevant sensory data from the remote robot
4. improve efficiency by reducing the need for supervision, thus allowing the operator to monitor multiple robots simultaneously
5. support the incremental evolution of telesystems to full autonomy”

## **Collaborative Control**

In human-robot interfaces, there must be a dialogue between the operator and the robot. The human should be able to express intent and interpret what the robot has done, while the robot should be able to provide contextual information and to ask the human for help when needed [Sheridan, 1992]. One approach to this type of interaction is collaborative control, a teleoperation model in which humans and robots work as peers to perform tasks [Sheridan, 1992].

Cooperative teleoperation tries to improve teleoperation by supplying expert assistance [Fong et al., 2001b]. Several robot control architectures have addressed the problem of mixing humans with robots. In a new control model, called collaborative [Fong et al., 2000] control, a human and a robot collaborate to perform tasks and to achieve goals. Instead of a supervisor dictating to a subordinate, the human and the robot engage in dialogue to exchange ideas and resolve differences. An important consequence is that the robot decides how to use the human's advice. With collaborative control, the robot has more freedom in execution. As a result, teleoperation is more robust and better able to accommodate varying levels of autonomy and interaction [Fong et al., 2001b].

Mixed-initiative interaction is an important aspect of effective multiagent collaboration to solve problems or perform tasks. Mixed-initiative refers to a flexible interaction strategy where each agent can contribute to the task what it does best. Furthermore, in the most general cases, the agents' roles are not determined in advance, but opportunistically negotiated between them as the problem is being solved [Allen, 1999]. At any one time, one agent might have the initiative—controlling the interaction—while the other works to assist it, contributing to the interaction as required. At other times, the roles are reversed, and at still other times the agents might be working independently, assisting

each other only when specifically asked. The agents dynamically adapt their interaction style to best address the problem at hand. The best way to view interaction between agents is as a dialogue, and thus mixed-initiative becomes a key property of effective dialogue [Allen, 1999].

In mixed-initiative interaction, the situation can be more complex. Because different agents might take the initiative at different times, an agent must be able to tell when it should appropriately start an interaction by taking the turn. Table 2.4 provides levels of mixed-initiative interaction based on [Allen, 1999].

Table 2.4. Levels of Mixed-Initiative Interaction

| MIXED-INITIATIVE LEVELS     | LEVELS                                                                     |
|-----------------------------|----------------------------------------------------------------------------|
| Unsolicited reporting       | Agent may notify others of critical information as it arises.              |
| Subdialogue initiation      | Agent may initiate subdialogues to clarify, and correct.                   |
| Fixed subtask initiative    | Agent takes initiative to solve predefined subtasks.                       |
| Negotiated mixed initiative | Agents coordinate and negotiate with other agents to determine initiative. |

Murphy [Murphy et al., 2001] describes a mixed-initiative system for urban search and rescue that was implemented on a team of heterogeneous robots. The mixed-initiative system was novel in that the robot took the initiative for perception rather than for navigation. The mixed-initiative control scheme employed a novel three-agent society organization that places an intelligent assistance agent as the middleware between the physically situated (remote robot) and cognitive (human) agents. The intelligent assistant

agent provides perceptual assistance, cueing the tele-operator to the possible presence of victims using a fusion of heat, motion, skin color, and color difference. Traditional mixed-initiative systems tend to focus on the blending of deliberation (planning) by the human and reaction (perception-action) by the robot. Murphy's approach to mixed-initiative systems divides tasks based on perception. The first perception task is victim detection of non-surface victims and requires distal perception. In these tasks, the human performs the explicit recognition and distal action while deciding what should be completed next.

This approach to mixed initiative control is novel for three reasons [Murphy et al., 2001],

1. It is an organization of intelligence.
2. It focuses on automating the perceptual tasks rather than the navigational tasks.
3. It uses collaborative teleoperation to improve navigation.

Murphy's [Murphy et al., 2001] approach contrasts with traditional mixed-initiative systems. The traditional approach consists of two agents and concentrates on relieving the operator of navigation tasks so that the operator can provide mission sensing. Murphy's [Murphy et al., 2001] work suggests the opposite: it is preferable for the physically situated agent (mobile robot) to provide mission sensors and the cognitive agent (human) to control navigation. Arguments for permitting the robot take the initiative in perceptual search including the following:

1. "The perceptual search for victims is more demanding than navigation in confined spaces. The amount of cognitive fatigue that the human experiences in this type of task is similar to that experienced by air-traffic controllers.

2. Most robot navigation is currently performed using range data extracted from sonar. Since the robot is operating in confined spaces, it is within the dead zone of the sonar range. The sonar transducers are also exposed to mud, water or dirt in this type of environment. Even if the robot were capable of semi-autonomous navigation, the sensors and thereby navigational autonomy would degrade over time.” [Murphy et al., 2001]

A middleware agent is needed to facilitate transactions between the cognitive and situated agents given the differences in cognitive ability, representation and contextual knowledge. The third agent is called the Intelligent Assistant Agent (IAA) and resides within the local workstation or wherever the user interface is generated [Murphy et al., 2001]. The IAA consists of a number of sub-agents such as the vision agent that takes the initiative in examining the perceptual data and displaying the results to the operator. The IAA fuses streams of concurrent sensor data in order to cue the operator to possible victims.

The operator communicates with the remote agents as needed and the IAA facilitates communication as well as handles the representation of data from the robots. The vision agent performed cueing and behavioral fusion based on the output of four concurrent detection algorithms: motion, skin color, color difference, and thermal region. The agents for the color camera were skin color detection, motion detection and a color difference algorithm.

### **Traded Control**

Traded control is another type of robot interaction that is closely related to collaborative control. Traded control is a situation in which a supervisor controls a robot during part of a task and the robot is autonomous during other portions of a task [Kortenkamp, 1997]. A significant problem in traded control situations is that the robot

does not know how the environment has been changed or what parts of the task have been accomplished when the human has been in control [Kortenkamp, 1997].

The goal of this work was to establish effective human-robot teams that accomplish complex tasks. Since the robot is a member of a human-robot team, it must be an equal partner with the human in performing those tasks. The software systems controlling such robots must allow for fluid trading of control among team members, whether they are humans or robots. This is the essence of mixed initiative interaction. Figure 2.9 is the progression of control from complete teleoperation to full autonomy.

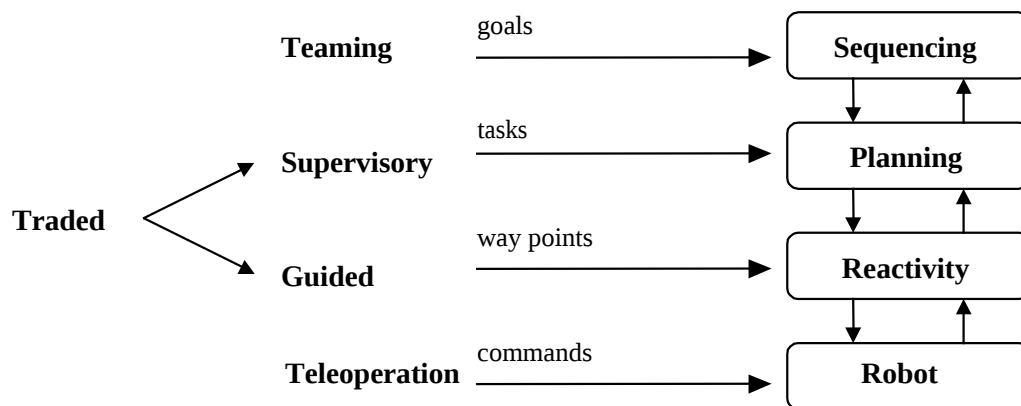


Figure 2.9. Teleoperation to Full Autonomy control Levels [Kortenkamp, 1997]

In teaming, robots and humans work as a team in which each member has full autonomy, but members communicate to accomplish complicated tasks. Interaction is at the planning level with goals given to the robot just as they are given to the other team members [Kortenkamp, 1997].

In supervisory control, robots work nearly autonomously, but a human is watching and can stop or correct the robot at any time. Interaction is at the task (sequencing) level, and the human has the opportunity to rearrange the robots' task plans or to stop the robot completely [Kortenkamp, 1997].

In traded control as defined by Kortenkamp, robots perform most tasks completely autonomously, but sometimes a human takes complete control to perform some difficult subtask or to extract the robot from a dangerous situation. Interaction is at the task (sequencing) level or through skills. Traded control is a mixture of supervisory and guided control [Kortenkamp, 1997].

In guided control, the human is always guiding the robot through a task although the robot has some autonomous capabilities, such as obstacle avoidance or grasping.

In teleoperation, the human is in complete control of the robot movements at all times. The robot has no autonomy. Interaction is with the mechanical robot servos and bypasses the architecture completely [Kortenkamp, 1997].

Effective traded control requires a robot system that can autonomously perform routine operations yet give control to a human to perform specialized difficult operations. The unique capabilities can be used when needed and not during tedious, repetitive and routine operations. Mixed initiative planning must not only plan for collaboration, but also re-plan in case of unanticipated sequencing.

Traded control allocates to the human the choice of an agent to perform a task. In semi-autonomous control, the robot can perform single primitive actions without human intervention. It is necessary to maintain the robot's awareness of the situation during human operations. In one approach to accomplish this, the robot monitors the teleoperated actions using sensor feedback events provided by the skill manager for autonomous execution. In

the second approach, the robot uses alternative sensing techniques to monitor the action. The user interface requirements for traded control are:

1. “The human and machine must exchange information on the machine's status, goals, beliefs, and intentions.
2. The human and machine must coordinate during joint or shared tasks.
3. The machine or human must update worldviews at task hand over.

[Kortenkamp, 1997]”

### **Applications**

There are several applications for human-machine interfaces including aviation and military operations. In aviation applications, one type of aviation display is the tunnel in the sky display. This display is an egocentric immersed viewpoint because the viewpoint in the display corresponds to the viewpoint of the observers as if they were immersed within the scene. Displays that contain an egocentric viewpoint preserve the visual relationships for tracking performance [Doherty et al., 2001]. In Wickens' [Wickens et al., 1997] research, three prototype displays are contrasted for aircraft navigation and tactical hazard awareness: a conventional 2-D coplanar display, an exocentric 3-D display, and an immersed 3-D display [Wickens et al., 1997]. The results of the study found that the immersed 3-D displays appear to be the most beneficial for guidance tasks and for navigational checking.

In military applications, interfaces are used in battle space visualization to enhance the commander's ability to understand the unfolding battle in order to make timely and informed tactical decisions [Barnes et al., 1998]. The intuitive understanding of the battle process includes visualization of possible end states and their consequences. The focus of the study by Barnes [1998] was to determine the behavioral link between different

representation techniques and the human's ability to better understand and make decisions about the battle process. A concern in military interface design is that it must provide ease of use and comfort while making the machine "transparent " and capable of sustaining the soldier for extended periods. The soldier should attend to the tactical situation and not the machine. The interface must provide perception, cognition and action [Barnes et al., 1998]. Evaluation of the battle space interface showed the immersive point of view characterizes virtual reality systems and may enable a soldier's eye view of the battlefield. A disadvantage of the immersive point of view is that it may cause a keyhole effect in that a more global perception of the terrain cannot be realized. The total immersion in the battle scene may improve situation awareness but the costs is that of missed information concerning other important battle parameters. Navigation is best supported by the immersive viewpoint, in part because this provides a frame of reference that is compatible with the view that the eyes normally have as the traveler navigates through real space [Banks et al., 1997].

This chapter has presented applications for human-machine interfaces in the disciplines of aeronautics and military operations. This chapter presented several types of interfaces including direct telemanipulation, sensor fusion, and virtual reality. Finally, the different types of control, such as teleoperation, supervisory, and collaborative control were discussed. The next chapter will highlight the basics of human-robot interfaces.

## **CHAPTER III**

### **EVALUATION OF HUMAN-MACHINE INTERFACES OVERVIEW**

This chapter is a literature review of metrics employed in the evaluation of human-machine interfaces, such as usability, mental workload, user satisfaction, and situational awareness. In order to evaluate the effectiveness and performance of an enhanced graphical user interface, it is necessary to define evaluation metrics and compare them to the original interface. This chapter presents interface attributes and methods for obtaining those attributes. This chapter also presents evaluation methods employed in Human Factors.

#### **Measured Attributes**

Several measures are used to evaluate human-computer interfaces. One method employs a subjective measure in which participants are asked to write down their impressions regarding a particular interface aspect. Quantitative performance measures are based on measured quantities such as response time and completion accuracy. The relative advantage of subjective measures is that they may address more general or "cognitive" issues [Sutcliffe, 1989].

#### **Examples of subjective measures are [Sutcliffe, 1989]:**

1. ease of use
2. ease of learning
3. preferences and reasons for preferences
4. predictions of behavior and reasons for predictions

The relative advantage of performance measures is that they are more objective and provide the precision necessary to fine-tune an interface while providing the ability to make more precise statements regarding the relative advantages of different approaches [Sutcliffe, 1989].

Examples of performance measures are [Sutcliffe, 1989]:

1. task completion time
2. menu navigation time
3. number of wrong menu choices
4. observation of frustration

### **Usability**

“Usability is defined as the user's ability to complete a task, efficiently and effectively, without undue stress and fatigue [Bevan, 1995a]”. The definition of usability is also the extent to which a user can use a product with satisfaction in a specified context. Usability is the capability of the software product to be understood, learned, used, and be attractive or appeal to the user [Bevan, 1995a]. Specifically, usability depends on who the participants are, as well as their ability to carry out their tasks and their goals. Effectiveness is the extent to which a user's task can be achieved. Efficiency is the amount of effort required to accomplish a task, and this may be measured by examining task execution time, error rates, as well as physical and/or mental workload [Bevan, 1995a]. Satisfaction is defined as the comfort and acceptability of use [Bevan, 1995a]. Usability can be evaluated by [Bevan, 1995a],

1. analysis of the product features

2. analysis of the interaction process
3. analysis of the effectiveness and efficiency

Table 3.1 lists some typical metrics for measuring product usability [Bevan, 1995a].

Table 3.1. Metrics for measuring usability

| Effectiveness Measures                                  | Efficiency Measures                  | Satisfaction Measures          |
|---------------------------------------------------------|--------------------------------------|--------------------------------|
| Percentage of goals achieved                            | Time to complete a task              | Rating scale for satisfaction  |
| Percentage of participants successfully completing task | Tasks completed per unit time        | Frequency of discretionary use |
| Average accuracy of completed tasks                     | Monetary cost of performing the task | Frequency of complaints        |

“User performance is measured by the extent to which the intended goals of use are achieved (effectiveness) and the resources such as time, money, and mental effort that have to be expended to achieve the intended goals (efficiency) [Bevan, 1995b]”. Satisfaction is measured by the extent to which the user finds the product acceptable. Bevan [1995b] defines the usability framework as show in Figure 3.1.

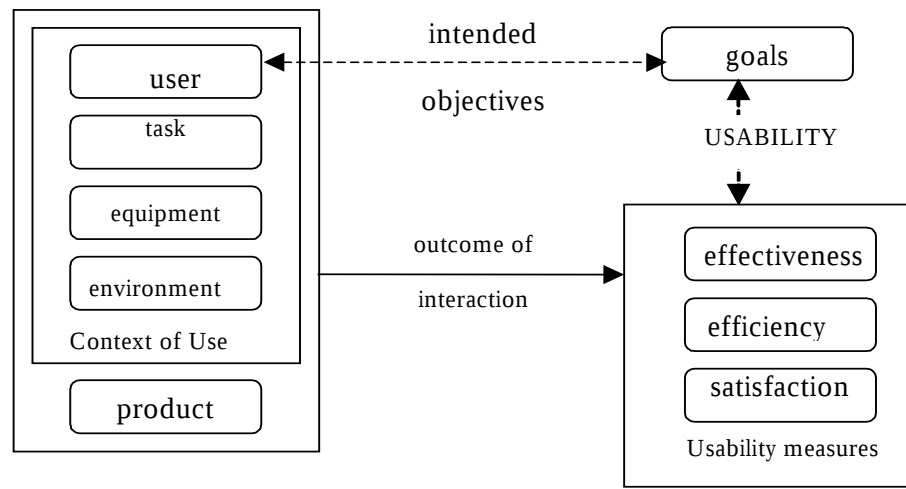


Figure 3.1. Usability Framework [Bevan, 1995a]

There are five primary attributes of usability. The five attributes are learnability, efficiency, memorability, errors, and satisfaction [Bevan, 1995b]. The system should be easy to learn and enable high user productivity. The system should also decrease the user's learning curve by reducing the amount of information that must be memorized. The system should have a lower error rate and enable easy error recovery. Finally, the system should be pleasant to use such that the user feels satisfied upon the task completion.

### **Mental Workload**

Mental workload relates to the mental effort required to perform tasks [Bevan, 1995a]. Mental workload is a useful measure when participants are required to expend excessive mental effort to achieve acceptable performance. Mental workload is particularly important in safety-critical applications. Adequate usability measures should include aspects of mental effort as well as performance [Bevan, 1995a]. Since mental workload is a

relationship between the amount of resources available as well as demanded in a task, this value can be changed by altering the resources available or demanded [Wickens et al., 1984].

Figure 2.10 shows the relationship between resource supply, demand, and task performance.

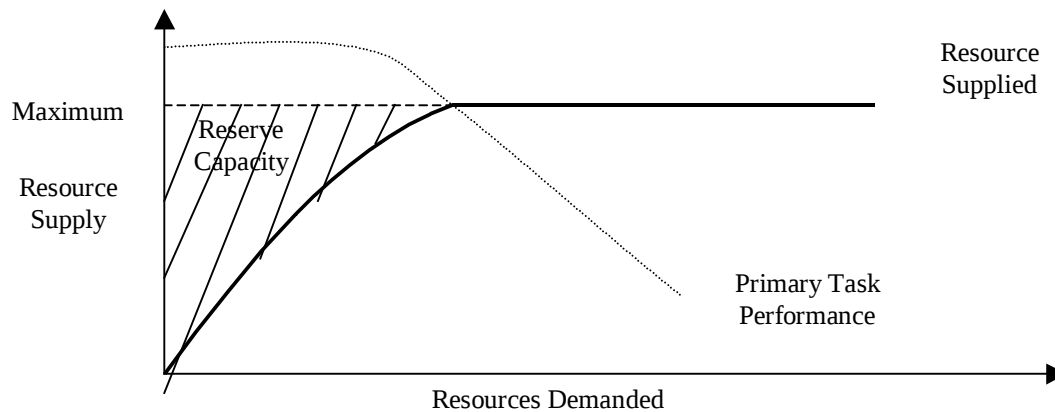


Figure 2.10. Relationship between resources and task performance [Wickens et al., 1984]

Mental workload can be measured by physiological parameters, dual task performance, or subjective workload assessments [Wickens et al., 1984]. Some of the benefits of understanding mental workload appear in designing an adaptive interface that adapts to increasing/decreasing workload. The workload and mental workload equations are [Burnett, 2001a]:

$$\text{workload} = \text{task completion time} / \text{time available to complete task}$$

$$\text{mental workload} = \frac{\text{required resources for task completion}}{\text{resources available for task completion}}$$

High mental workload levels can lead to additional stress and decreased operator accuracy.

One proposed solution for reducing mental workload is to use automation [Burnett, 2001a].

One problem found with subjective workload assessment techniques is that they do not accurately reflect objective workload. For example, a user may rate a task as not requiring much effort, but use near maximum information processing resources in order to complete the task. Studies show that participants do not tolerate high workload and adapt their behaviors and actions in order to adjust to the workload level.

Mental workload consists of objective factors such as number of tasks, urgency, and cost of non-completion of the task on time or correctly, as well as a range of subjective factors and environmental variables. Performance tends to decrease steeply when workload becomes too high, while through the normal range of workload little change is noticed. It is worth noting that if operators have too little to do, performance. Figure 3.2 indicates the relationship between mental workload level and performance.

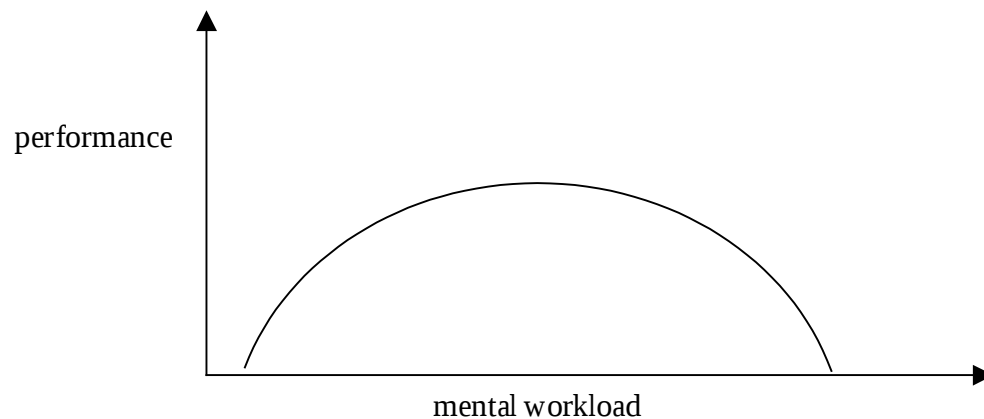


Figure 3.2. Performance vs. Mental Workload [Burnett, 2001a]

## Situational Awareness

Situational awareness (SA) is the knowledge of what is going on around the human operator or the robot [Endsley et al., 2000]. Situational awareness incorporates an operator's understanding of an entire situation so that they can form a basis for decision making. SA and mental workload are independent variables but increased mental workload may have a negative effect on SA in certain situations. Table 3.2 demonstrates a relationship between SA and mental workload.

Table 3.2. Relationship between SA and Workload [Endsley, 1995]

| Situational Awareness | Workload | Description                                                                                                                                                                                                                                |
|-----------------------|----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| LOW                   | LOW      | The operator may have little idea of what is going on and is not actively working to find out because of inattentiveness, vigilance problems or low motivation                                                                             |
| LOW                   | HIGH     | If there is a great number of tasks, SA may suffer because the operator can attend only to a subset of information or may be working to actively achieve SA, and have an erroneous or incomplete perception and integration of information |
| HIGH                  | LOW      | Information can be presented in a manner that is easy to process (an ideal state)                                                                                                                                                          |
| HIGH                  | HIGH     | Operator is working hard but is successful in achieving accurate and complete picture of the situation                                                                                                                                     |

There are three levels of situational awareness [Endsley et al., 2000]: perception, comprehension, and prediction. Level 1 is the perception of the status, attributes, and dynamics of relevant elements of an environment. Level 2 is a comprehension of the situation based upon the elements perceived in Level 1. The elements of Level 1 are synthesized to form patterns and influence decisions. Level 3 is the ability to predict future

actions of perceived elements based upon the comprehension of the situation in Level 2 [Endsley et al., 2000]. Figure 3.3 illustrates the definition of the term “situation” by embedding it in the action-perception loop of a situation-oriented behavior-based robot. The actions of the robot change the state of the environmental model, and the robot’s sensors perceive some of these changes. The human operator must assess the robot’s situation and select an appropriate goal or behavior. The role of the human operator is to define external goals and to control behavior selection via a human-robot interface [Graefe, 1998].

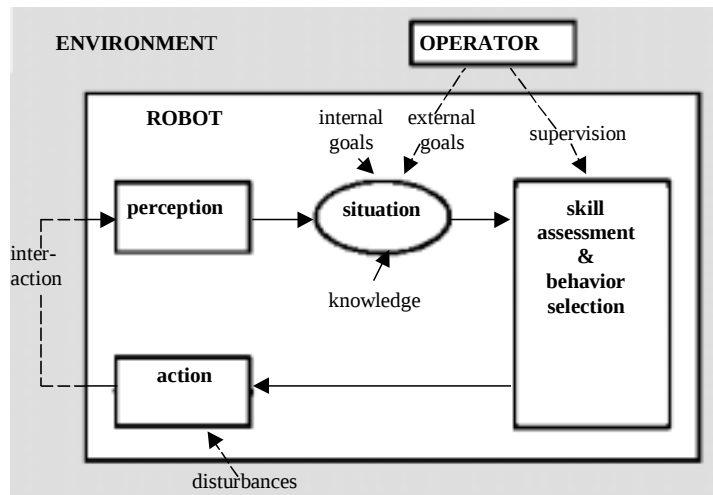


Figure 3.3. Situation in the action-perception loop [Graefe, 1998]

SA is temporal in nature therefore an operator’s awareness of an environment is constantly changing [Endsley, 1995]. The primary purpose of improving an operator’s SA is to provide the operator with the necessary information to diagnose and resolve unexpected events at the remote site. There are several types of SA errors based upon the level when it occurs. Level 1 errors are when a person fails to perceive information that is important to an assigned task. This is closely tied to the belief that the value of some important variable is

different from what it actually is. Level 2 errors are the result of the user's inability to comprehend perceived data with reference to user goals. The user cannot determine the relevance of elements important to those goals. Level 3 errors make it difficult for the user to project future events based upon the elements comprehended in level 2. Some possible measures that may determine if a user will be good at SA are spatial ability, perceptual ability, logical ability, personality factors, memory, and attention ability [Endsley, 1995].

## **Evaluation Methods**

### **Heuristic Evaluation**

A Heuristic Evaluation involves a group of interface participants examining an interface in order to identify violations of interface design principles [Prothero, 1994]. Heuristic Evaluation is a process of applying 'golden rules' of effective interface design to a target system. The evaluation process involves walking through the interface and assessing which interface aspects are in agreement with these rules. Empirical studies of heuristic evaluations show that experts determine approximately 50% of the usability problems [Wild et al., 2000]. Most heuristic evaluations are performed by at least three or more participants that independently examine the interface and report any problems. This method shows graceful degradation when used with non-HCI specialists. This method can also be performed in conjunction with other evaluation methods. The main advantage of Heuristic Evaluation is that it can be applied quickly and cost effectively [Wild et al., 2000].

## **Cognitive Walkthrough**

The Cognitive Walkthrough is a method of usability evaluation that uses a theory of learning by exploration [Wild et al., 2000]. The input to a Cognitive Walkthrough is a detailed description of an interface, a task scenario, and the explicit assumptions regarding the user population and the context of use [Wild et al., 2000]. Cognitive Walkthroughs consist of participants answering a set of questions regarding the decisions that an interface user must make and rating the likelihood that the user will make an incorrect choice [Prothero, 1994].

## **MUSiC**

The MUSiC methods were specifically developed by the European MUSiC (Metrics for Usability Standards in Computing) project to provide valid and reliable means of specifying and measuring usability [Macleod et al., 1997]. MUSiC provides diagnostic feedback that enables the design to be modified in order to improve usability [Macleod et al., 1997]. The current MUSiC definition of usability is [Macleod et al., 1997]:

“the ease of use and acceptability of a system or product for a particular class of participants carrying out specific tasks in a specific environment; where ‘ease of use’ affects user performance and satisfaction, and ‘acceptability’ affects whether or not the product is used”

## **SUMI**

The assessment of a computer system's usability should involve measuring not only aspects of participants' performance, but also how participants feel about the system they are using [Bevan et al., 1997]. To measure user satisfaction and hence assess user perceived

software quality, the University College Cork has developed the Software Usability Measurement Inventory (SUMI) as part of MUSiC [Bevan et al., 1997]. SUMI is an internationally standardized 50-item questionnaire, available in seven languages. It takes approximately ten minutes to complete and contains statements for the evaluator to rate such as [Bevan et al., 1997],

- "Using this software is frustrating
- Learning how to use new functions is difficult

SUMI results have been shown to be reliable and to discriminate between different types of software products in a valid manner. SUMI provides an overall assessment and a Usability Profile that breaks the overall assessment down into five sub-scales: affect, efficiency, helpfulness, control, and learnability [Bevan et al., 1997].

### **NASA-TLX**

The NASA-Task Load Index (TLX) is an internationally used and acknowledged workload assessment technique [Hart et al., 1998]. NASA TLX is used to understand a participants' perceived workload. The Task load Index is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six subscales. Three subscales relate to the demands imposed on the participants in terms of [Hart et al., 1998]:

1. the amount of mental and perceptual activity required by the task
2. the amount of physical activity required by the task
3. the time pressure felt due to the task

Another three subscales relate to the interaction of an individual with the task [Hart et al., 1998]:

1. the individual's perception of the degree of success
2. the degree of effort an individual invested
3. the amount of insecurity, discouragement, irritation, and stress

Table 3.3 is a description of the six NASA-TLX subscales.

Table 3.3. Subscales of the NASA-TLX [Hart et al., 1998]

| TITLE             | DESCRIPTION                                                                                                                                                                                      |
|-------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Mental Demand     | How much mental and perceptual activity was required?<br>Was the task easy or demanding, simple or complex, exacting or forgiving?                                                               |
| Physical Demand   | How much physical activity was required?<br>Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?                                                             |
| Temporal Demand   | How much time pressure did you feel due to the rate or pace at which the task or task elements occurred?<br>Was the pace slow and leisurely or rapid frantic?                                    |
| Performance       | How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)?<br>How satisfied were you with your performance in accomplishing these goals? |
| Effort            | How hard did you have to work (mentally and physically) to accomplish your level of performance?                                                                                                 |
| Frustration Level | How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?                                             |

After administration of the NASA-TLX, Cha [Cha et al., 1997] discovered that participants had problems rating perceived workload for many of the scales. The scales consisted of technical, vague and unfamiliar words for the common user. Therefore, the

RNASA-TLX was developed for a driving application and the scales were modified to reflect the direct application to specific tasks [Cha et al, 1997]. The highest ranked factors of mental workload for the RNASA-TLX were determined to be visual demand then mental demand, difficulty in driving, temporal demand, difficulty in understanding information, and the auditory demands. Table 3.4 lists the rating scales for the RNASA-TLX.

Table 3.4. RNASA-TLX rating scales [Cha et al., 1997]

| TITLE                                   | DESCRIPTION                                                                                                                                                     |
|-----------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Mental Demand                           | How much mental attention was needed during driving when using the IVNS?                                                                                        |
| Visual Demand                           | How much visual activity was required while driving when using an IVNS to recognize the information from an IVNS or other external information sources?         |
| Auditory Demand                         | How much auditory activity was required during driving when using an IVNS to recognize or hear the information presented from an IVNS or other auditory source? |
| Temporal Demand                         | How much time pressure was required due to rate or pace as the task elements occurred during driving using an IVNS?                                             |
| Difficulty in Driving                   | How hard was driving when using an IVNS compared with other in-vehicle control equipment or optional devices?                                                   |
| Difficulty in understanding information | How hard was it to understand information presented from an IVNS?                                                                                               |

In both the RNASA-TLX and the NASA-TLX, each rating scale value is weighed by the amount the user feels this aspect affects overall mental workload. After comparison of several subjective workload assessments, it was determined that the TLX method provides more consistent scores among people doing the same task. Appendix A contains an example of a NASA-TLX development.

### **Subjective Workload Assessment Technique (SWAT)**

SWAT is a subjective rating that uses three levels- low, medium, and high to rate particular tasks. The workload is rated over the three scales: time load, mental effort, and psychological stress [Wierwille et al., 1993]. The first step in SWAT development is for the user to rate 27 permutations of these three rating scales. After the user rates the task in terms of each rating scale, the value is weighted by the numerical value from step one[Wierwille et al., 1993; Cha et al., 2001].

Since SWAT and NASA-TLX are multidimensional, it is possible to obtain workload measures across sub scales. In a comparison of NASA-TLX and SWAT, NASA-TLX consistently shows higher workloads than SWAT [Wierwille et al., 1993]. SWAT has a greater potential for identifying workload factors such as cognitive mechanisms affecting mental workload judgments [Wierwille et al., 1993].

### **Multiple Resource Questionnaire (MRQ)**

The Multiple Resource Questionnaire is a 17-item test administered to participants in order to measure subjective workload [Boles et al., 2001a]. The MRQ is based upon multiple resource theory and provides high diagnosticity of certain workload resources. One disadvantage of SWAT and NASA-TLX is that participants have to sort through a large amount of workload dimensions and sort them [Boles et al., 2001b]. The sorting procedure is very monotonous and requires extensive user contribution before the rating process begins. MRQ measures workload based upon the multiple resources technique and does not require sorting. Studies show MRQ has proven to be as reliable as SWAT and NASA-TLX without the additional sorting procedure. Some of the items included in the MRQ are the auditory emotional process, short-term memory process, spatial concentrative process, and

spatial quantitative process [Boles et al., 2001b]. Some of these measures have been used in this study and are presented in Appendix B.

### **Spatial Reasoning**

The Vandenberg Mental Rotations test contains 20 items in five sets of four items [Vandenberg et al., 1979; Shepard et al., 1971]. Each item consists of a criterion figure, two correct alternatives, and two incorrect alternatives or “distractors”. Correct alternatives are always identical to the criterion in structure but are shown in a rotated position. The distractors are rotated mirror-images of the criterion or rotated images of one or two criteria. For scoring, a line is counted as correct if both choices are correct. This method of scoring eliminates the need to correct for guessing [Vandenberg et al., 1979]. Figure 3.4 provides some sample questions from the Vandenberg mental rotation test. The entire test can be found in Appendix C.

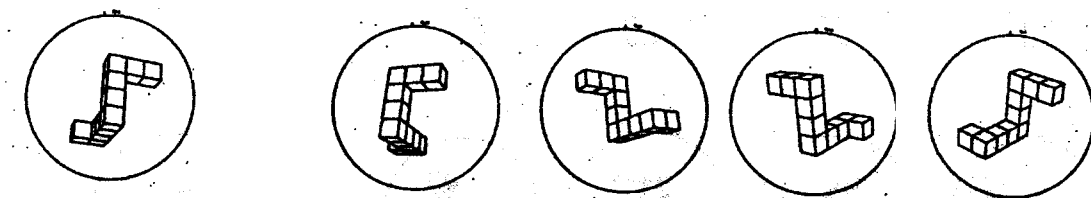


Figure 3.4. Mental Rotation Sample Questions [Vandenberg et al., 1979]

### **Contextual Inquiry**

Contextual Inquiry (CI) is a structured interviewing method for evaluating the design of interactive systems in the context of the work being performed [Holtzblatt et al., 1996].

Since CI is subjective, it is most appropriate for qualitative system assessment rather than for performance measurements. CI is usually performed during the design process for a user interface.

An interviewer performs CI by observing participants while they work and asking questions as they perform tasks in order to understand their motivation and strategy. Typical questions are: “What are you doing now?” , “Is that what you expected to happen?” , “What do you especially like or dislike about this tool?”[Holtzblatt et al., 1996]. Through conversation and discussion, the interviewer and the user develop a shared understanding of the work. Thus, CI supports system development by providing a mechanism to help identify and articulate the task [Holtzblatt et al., 1996].

CI is based on three principles: context, partnership, and focus [Holtzblatt et al., 1996]. Context implies participants describe their work as they perform tasks in their normal working environment. Partnership is the concept that the user should share in guiding the design process. The key to a partnership is maintaining a conversation that permits the user and the interviewer to create a shared understanding regarding the work experience. Focus describes the objectives that the study is attempting to achieve. Focusing on specific goals guides what is attended to or ignored, what questions are asked, and what is probed further [Holtzblatt et al., 1996].

A fundamental problem in CI is how the interviewer can encourage the user to “open up” and provide key information. One approach is the apprenticeship model [Holtzblatt et al., 1996]. With this model, the interviewer acts as an apprentice and asks the user to teach the apprentice how to use the system and complete the work. The apprenticeship model encourages participants to shape and guide the conversation. It also

helps ground the conversation on concrete details, what procedures are needed for work, and where the problems are located [Holtzblatt et al., 1996].

### **Cooperative Evaluation**

Cooperative Evaluation is a variant of the think-aloud observational technique [Holtzblatt et al., 1996; Wild et al., 2000]. While the participants perform tasks, they are asked to think aloud regarding their experience and problems [Wild et al., 2000]. This approach is low in resource costs and can be employed by non-HCI experts to generate ecologically grounded usability problems. Furthermore, the Cooperative Evaluation method demands an initial task decomposition and requires developers to work with participants [Wild et al., 2000]. This method does not work well because it is difficult for participants to remember to articulate their thoughts and problems.

In a modification of the cooperative evaluation, the user thinks aloud both before and after the task [Gediga et al., 2001]. This is sometimes referred to as the pre-event and post-event procedure. This technique would be useful when the user needs to concentrate on the task and question probing would cause a distraction. Participants' comments are recorded while they view the video record of the task execution. At times, this method has been criticized because during post-event protocols, participants might rationalize their own actions [Gediga et al., 2001]. The empirical comparison of post-event protocols with event protocols shows that the information provided by post-event protocols are of higher quality, although the amount of information is reduced in comparison to event protocols [Gediga et al., 2001].

Another variation on the cooperative evaluation is the video confrontation. The video confrontation is based upon a post-event thinking-aloud technique. The experimenter

selects certain segments of the video and interviews the participant about these. Since this method focuses on salient points, the protocol is much easier to analyze than the general think-aloud [Gediga et al., 2001]. The steps for the cooperation evaluation method are recruit participants, prepare tasks, as well as interact, record, and summarize observations. The video confrontation method of Cooperative Evaluation was used as part of the interface evaluation for this research.

### **Situational Awareness Rating Technique (SART)**

Selcon [Selcon et al., 1989] developed SART in order to accurately measure pilots' situational awareness (SA) by taking into account mental workload factors [Selcon et al., 1989]. This study stated that the common elements of SA are pilot knowledge, understanding of goals, and tactical awareness. SART is a ten-dimensional scale with three major groupings of the ten items on the scale. The three groupings are demand on attentional resources, supply of attentional resources, and understanding [Selcon et al., 1989]. Table 3.5 shows the ten-dimensional SART scale.

In conclusion, there are several parameters of a human-robot interfaces that can be measured from the human factors engineering perspective. These measured attributes include usability, mental workload, and situational awareness. Evaluation methods for these attributes are heuristic evaluation, cognitive walkthrough, MUSiC, SUMI, NASA-TLX, MRQ, SWAT, Spatial Reasoning, Contextual Inquiry, Cooperative Evaluation, and Usability Questionnaires. The methods used to evaluate the enhanced user interface were the NASA-TLX, MRQ, and Usability Questionnaires. The final method was a variation on the contextual inquiry where the participant was asked about their thoughts during the task while watching a videotape of the task execution.

Table 3.5. Ten-dimensional SART Scale [Selcon et al., 1989]

| DEMAND                   | LOW.....HIGH |   |   |   |   |   |   |
|--------------------------|--------------|---|---|---|---|---|---|
| Instability of situation | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| Variability of situation | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| Complexity of situation  | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| SUPPLY                   |              |   |   |   |   |   |   |
| Arousal                  | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| Spare mental capacity    | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| Concentration            | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| Division of attention    | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| UNDERSTANDING            |              |   |   |   |   |   |   |
| Information Quantity     | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| Information Quality      | 1            | 2 | 3 | 4 | 5 | 6 | 7 |
| Familiarity              | 1            | 2 | 3 | 4 | 5 | 6 | 7 |

## **CHAPTER IV**

### **SENSORY EGOSPHERE**

This chapter discusses the concept of a Sensory EgoSphere (SES) that is employed to enhance a graphical user interface to a mobile robot. The chapter reviews the origins of the Sensory EgoSphere along with the migration to the current use, as defined in this research. James Albus [Albus, 1991] first proposed a Sensor EgoSphere in 1991. He defined the SES as a dense spherical coordinate system with the self (ego) at the origin. This concept was proposed as part of his outline for the theory of machine intelligence. In order to discuss the origins of the SES, it is necessary to review Albus' intelligence theory.

Albus defined intelligence as that which produces successful behavior [Albus, 1991]. He proposed a model that integrated knowledge from research in both natural and artificial systems. The ultimate goal was the development of a general theory of intelligence that encompasses both biological and machine instantiations [Albus, 1991]. At a minimum, Albus stated that intelligence requires the ability to sense the environment, to make decisions, and to control action [Albus, 1991]. Intelligence should improve an individual's ability to act effectively and wisely choose between alternative behaviors. There are four system elements to intelligence: sensory processing, environment modeling, behavior generation, and value judgment [Albus, 1991]. This review focuses on sensory processing and environment modeling for their applications to the SES. Sensors input information into an intelligent system and actuators output information from an intelligent system. The sensor provides input into a sensory processing system. Perception is the result of sensory processing. Sensory processing integrates similarities and differences between observations

and expectations over time and space in order to detect events. Sensory processing also recognizes features, objects, and relationships in the world [Albus, 1991].

Sensory perception is the transformation of data from sensors into meaningful and useful world representations. Sensory perception accepts input data from sensors that measure external world states as well as the internal system states. Perception scales and filters data. Sensory perception also computes observed features and attributes, while comparing observations with expectations generated from internal models [Albus, 1991].

The environment model is an intelligent system's best estimate of the environment state. The environment model includes a knowledge database regarding the world in addition to a database management system that stores and retrieves information [Albus, 1991]. Environment modeling uses sensory input to construct, update, and maintain a knowledge database. This is the function of the short and long-term memory [Albus, 1996].

### **Sensor EgoSphere**

In order to visualize an environment model, a map is typically required. A map is defined as a two-dimensional database that defines a mesh or grid on a surface [Albus, 1991]. The surface represented by a map may or may not be flat. For example, a map may be defined over a surface that is draped, or even wrapped around, a three dimensional volume. There are three general types of map coordinate systems that are relevant to an intelligent system: world coordinates, object coordinates, and egospheres [Albus, 1991]. World coordinate maps are typically 2-D arrays that represent projections of the earth's surface along the local vertical. World coordinates are often expressed in a Cartesian frame, and referenced to a point in the world [Albus, 1991]. Object coordinates are defined with

respect to features in an object. The origin may be defined as the center of gravity. This discussion focuses on the latter, egospheres.

Egospheres are a two-dimensional spherical surface that is a world map as seen by an observer at the sphere center [Albus, 1991]. Visible points on regions or objects in the world are projected on the egosphere. The projection is located where the line of sight from a sensor at the center of the egosphere to the points in the world intersect the surface of the sphere. Albus [Albus, 1991] states that the egosphere is the most intuitive of all coordinate systems. Each of us resides at the origin of our own egosphere. Everything that humans observe can be described as being located at some azimuth elevation and range measured from the center of our ego. To the observer at the center, the world is seen as if through a transparent sphere. Each observed point in the world appears on the egosphere at a location defined by that point's azimuth and elevation.

Objects may be represented on the egosphere by icons, and each object may have in its database frame a trace, trajectory, or position on the egosphere over some time interval. An ego motion occurs when the self-object moves through the world. The egosphere moves relative to world coordinates, and points on the egocentric map flow across their surfaces. Ego motion may involve translation, rotation, or both in a stationary world, or a world containing moving objects. If ego motion is known, the range to all stationary points in the world can be computed from the observed image flow. Once the range to a stationary point in the world is known, its pixel motion on the egosphere can be predicted from knowledge of ego motion. For moving points, prediction of pixel motion on the egosphere requires additional knowledge of object motion.

A number of different egosphere coordinate frames are useful for representing the world. These include the sensor egosphere, head egosphere, body egosphere, inertial

egosphere and, velocity egosphere [Albus, 2001]. A sensor egosphere is an egosphere in which the horizontal axis of the sensor array defines the egosphere equator and hence the pole. The center pixel in the sensor array defines zero azimuth at the equator. Sensor EgoSphere coordinates are defined by the sensor position and orientation. The coordinates move as the sensor moves. Figure 4.1 is a graphic of the Sensor EgoSphere as defined by Albus.

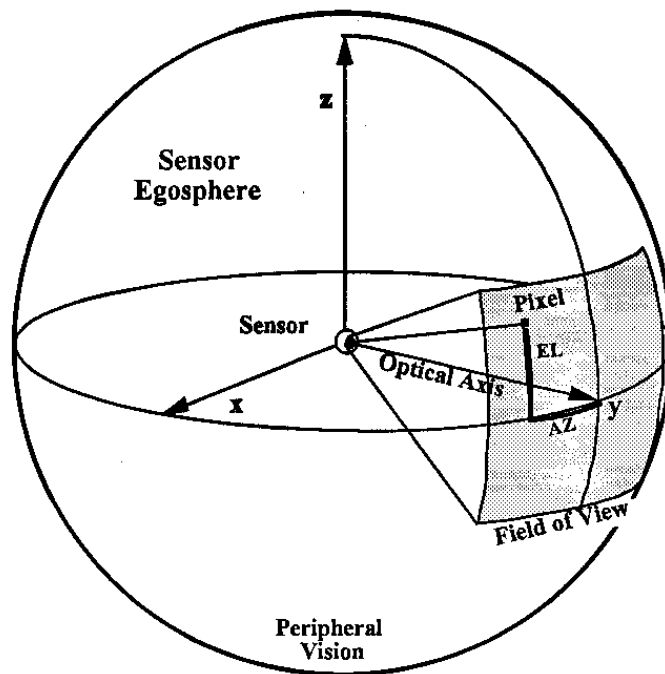


Figure 4.1. Sensor EgoSphere for a Camera [Albus, 2001]

Real-time sensory data can be employed to build world maps and provide an environment model to a human user. High-resolution dynamic information must be generated from real-time sensory data. Cameras, laser, and stereo systems can provide range information in egosphere coordinates. This information can be employed to build local

terrain maps in real time as well as represent moving objects. When the position and orientation of a camera egosphere is known, local maps generated from camera data can be registered with a priori maps. The result enables landmark recognition and provides the required information to permit supervisory control, path planning, obstacle avoidance, and task decomposition [Albus, 2001]. The foundations of the development of the Sensor Egosphere have lead to the implementation of the Sensory EgoSphere in the Vanderbilt University Intelligent Robotics Laboratory.

### **Sensory EgoSphere**

The Sensory EgoSphere (SES) developed in the Intelligent Robotics Laboratory (IRL) is defined differently than that used by Albus. This Sensory EgoSphere is a biologically inspired short-term memory structure [Peters et al., 2001a]. The SES is implemented as a variable density virtual geodesic dome upon which sensory data from the area surrounding the robot is developed. The nodes are links to data structures and files in a database. The SES database is indexed by azimuth and elevation. The database is also searchable by location and content [Peters et al., 2001].

The SES is a relatively simple, computational database that embodies a subset of the natural functionality. The SES is a quasi-uniform triangular tessellation of a sphere into a polyhedron. It is the optimal solution to the problem of how to cover a sphere with the least number of partially overlapping circles of the same radius. The triangles connect at vertices forming twelve pentagons and a variable number of hexagons. The pentagons are evenly distributed so that the node at the center of one is connected to the centers of the five others by  $N$  vertices, where  $N$  is called the frequency of the dome. The number of vertices on the polyhedron is a function of the frequency. A frequency one dome is an

icosahedron that has twelve vertices, each of which connects with five neighbors. To be useful as a sensory data structure, the tessellation frequency should be determined by the resolution of the various sensors on the robot [Peters et al., 2001a].

### **Geodesic Dome Topology**

The SES definition, in this work, is a two-dimensional spherical data structure, centered on the robot coordinate frame. The SES is a sparse environmental map containing pointers to object or event descriptors that have been recently detected by the robot. As the robot operates in the environment, both external and internal events stimulate the robot's sensors. After the stimulus, the associated sensory processing module writes its output data to the SES at the node closest to the direction from which the stimulus arrived. Sensory data of different modalities coming from similar directions at similar times register close to each other on the SES [Peters et al., 2001a].

Given that sensors on a robot are discrete, there was no advantage in creating the SES as a continuous structure. Also the computational complexity of SES increases with its size and depends on the density of the tessellation. A virtual geodesic dome provides a uniform tessellation of vertices such that each vertex is equidistant to six neighbors.

The SES is a multiply linked list of pointers to data structures. There is one pointer for each vertex on the dome. Each pointer record has seven links, one to each of its six neighbors and one to a tagged-format data structure. A tag indicates the specific sensory data stored at that vertex. The time stamp indicates when the data arrived [Peters et al., 2001a].

The SES used in this study had an ATRV-JR robot as the ego center. The SES has also been developed for a Pioneer 2-AT robot. The ATRV-JR robot had two cameras, sonar

and laser sensors. The camera head is the center of the geodesic dome. Since most robots do not have sensory data that covers 360 degrees, the SES is not a complete geodesic dome. The SES is restricted to only those vertices that fall within the robot's sensory field. The camera is mounted on a pan-tilt head, therefore imagery or image features can be stored at the vertex closest to the camera direction. The sonar and laser extract data only around the robot equator so this data is posted to the SES equator [Peters et al., 2001a].

### **Short-term Memory**

The SES is a short-term memory structure. The SES records and recalls events that are localized in space or time. As a short-term memory, SES is useful in maintaining an object inventory in the robot's locale for subsequent manipulation or other actions. The spatial layout of the SES maintains the relationships between objects so that the robot knows where it is in relationship to the objects around it [Peters et al., 2001].

Short-term memory (STM) differs from immediate experience in that it persists after the stimulus is removed [Albus, 1996]. STM also differs from immediate experience in that it contains only symbolic representations. STM is dynamic unlike long-term memory and contains both symbolic and iconic representations of attention entities [Albus, 1996]. STM retains information by recirculation or rehearsal. If this recirculation is interrupted, or overwritten with new information, the previously stored information in short term memory is lost. STM provides a buffer between immediate experience and long-term memory. If entities and events detected in immediate experience are determined to be important, they can be transferred into the long-term memory [Albus, 2001]. Otherwise, such entities and events can be discarded or overwritten by subsequent inputs.

The concept of an SES was inspired by the Hippocampus, a structure common to mammalian brains [Peters et al., 2001a]. This is the mammal's primary short-term memory structure, and all cortical sensory processing modules communicate with it. While awake, the human hippocampus stores all incoming sensory information and associates the sensory responses to events that occur relative to each other in space and time [Peters et al., 2001a]. While asleep, the hippocampus translates this information from the short-term to long-term memory. This process is akin to the SES sensing the data as the short-term memory that is dynamic and changes when overwritten by new sensory instances at the same node [Peters et al., 2001a].

The rapid growth in computing power and concurrent decline in cost has dramatically increased the potential of robots to interact naturally with the world. Computers can process a sufficient quantity of sensory data quickly enough to permit a robot to adapt to a natural, unstructured environment. This adaptability requires the robot's sensory system to be properly organized while being appropriately coupled to the robot's actions. This adaptability also requires that the robot's actions in response to specific sensory input be changeable by the robot itself [Koku et al., 1999].

A sensory data set at a specific SES location can be stored as an object with an aging timer. Objects at a specific SES location can be deleted from the sphere after a specific period dependent upon the data type. The arrival of up-to-date sensory information can be employed to overwrite the older information at the same location. Methods for rapidly validating the currently presented data as well as the current state of the world are essential. This requirement implies a need for data specific descriptors [Koku et al., 1999].

Sensors and low-level sensory processing can provide an extensive set of spatial features at any point on the SES. Examples of such features are light intensity, hue, color, saturation, motion direction, and speed [Koku et al., 1999].

Structurally, the SES is a multi-resolution (multi-layered) database. Functionally, SES is a multi-user accessible database [Koku et al., 1999]. SES can be interpreted as an associative memory, where the association is through proximity [Koku et al., 1999]. Visible objects that have been recognized can be labeled on the SES, and the labels may be used as search keys. For example, if the robot is searching for a coke can in order to serve a beverage to a person, it would search for the words “coke can” as a query to the SES database. This query would return the most recent location where a coke can was spotted on the egosphere and narrow the visual search space significantly. SES provides a natural sense of spatial coherence and continuity so that imprecise interactions with the robot can become more specific [Koku et al., 1999].

### **Data Storage/Retrieval**

Sensory processing modules (SPM) write information to the SES. SPMs transmit a location, a tag, a time, and the pointer for the sensed data to the SES agent. The SES agent locates the vertex closest to the given location, then writes the tag and associated data in the vertex record, potentially overwriting any existent tag record with the same name. The SES also searches for the vertex or vertices that contain a given tag [Peters et al., 2001a].

The SES may also contain links to long-term memory (LTM) data structures [Peters et al., 2001a]. While the robot is stationary, the data the robot senses is added to the SES. If the sensed object is also stationary, data displayed on the SES will not move. To correct registration of moving objects requires object tracking. If the robot moves, the location of

the information display on the SES also moves as a function of the robot's heading and velocity [Peters et al., 2001a].

Sensory processing modules (SPM) transmit information to the SES through a software agent called the SES manager [Peters et al., 2001]. The SES manager interfaces to a standard database such as Microsoft Access<sup>TM</sup> or Standard Query Language (SQL). The SES manager determines the vertex closest to the given location and writes the tag and associated information in the database record associated with the node. Other agents such as those performing data analysis or data display can read from or display any given vertex on the SES. The manager also searches for the vertex or vertices that contain a given tag. The fixed number of nodes limits the search paths as the data displayed on the sphere increases. The various agents that display to the SES can also mark the data with a relevance or saliency metric [Peters et al., 2001]. The relevance or saliency metric estimates the importance of the information with respect to the current task or the robot's welfare. The SES agent maintains a list of active nodes, in other words, the vertices containing data.

The SES database has been implemented on a humanoid robot using a Microsoft Access<sup>TM</sup> database [Peters et al., 2001]. Visual Basic 6.0<sup>TM</sup> manages communications between the database and other system components. The database consists of a single table containing registered information. A software agent requiring access to the SES communicates with the SES manager agent. The SES manager agent relays the request to the database. The four request types include post data, retrieve data using data name, retrieve data using data type, and retrieve data using spatial location [Peters et al., 2001].

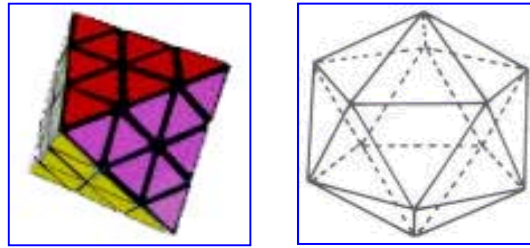
The post function compiles the relevant data from the requesting agent and registers the data in the database at the correct node location. Relevant data includes data name, data type, and the tessellation frequency at which the data should be registered. The node angles

are determined by the SES manager from a spatial direction and are included in the request. Visual information refers to the camera head pan and tilt angles when the image was captured. A request to retrieve data by data type results in all of the same data type being returned. A request to retrieve data by location returns all the data located at a specific node location.

### **Graphical Representation**

There are presently two graphical representations of the Sensory EgoSphere. The first is a geodesic dome using an octahedron as the basic shape. The octahedron SES was originally implemented using VRML. It was found that VRML exhibited a large degradation in performance with increased tessellation and texture mapping. The SES was re-implemented using OpenGL®. OpenGL® is a powerful graphics platform and can be programmed using Visual Basic™. The OpenGL® implementation is more compatible with the programming environment employed in the IRL's agent-based software architecture, Intelligent Machine Architecture (IMA) [Pack, 1998]. IMA permits the concurrent execution of software agents across multiple machines while facilitating extensive inter-agent communication.

The octahedron based tessellated dome is used with the ATRV-JR and Pioneer-2 AT mobile robots. The octahedron-based dome is shown in Figure 4.2a and the icosahedron based tessellated dome used with the IRL's Humanoid robot, ISAC is depicted in Figure 4.2b.



a.

b.

Figure 4.2. Octahedron- and Icosahedron- Based Tessellated Dome

Figure 4.3 illustrates the relative pose of the mobile robot inside the SES.

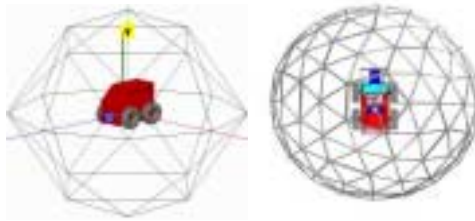


Figure 4.3. Relative position of Robot to SES [Johnson, 2002]

There are several sensor modalities available with the mobile robots. Sonar and laser are only effective in the mobile robot's equatorial plane; therefore, the resulting data is restricted to the vertices near the dome's equator. In order to simplify the display, the user has the option of displaying the sonar and laser data at the actual projection point. Figure 4.4 is a depiction of all the possible sonar and laser representations.

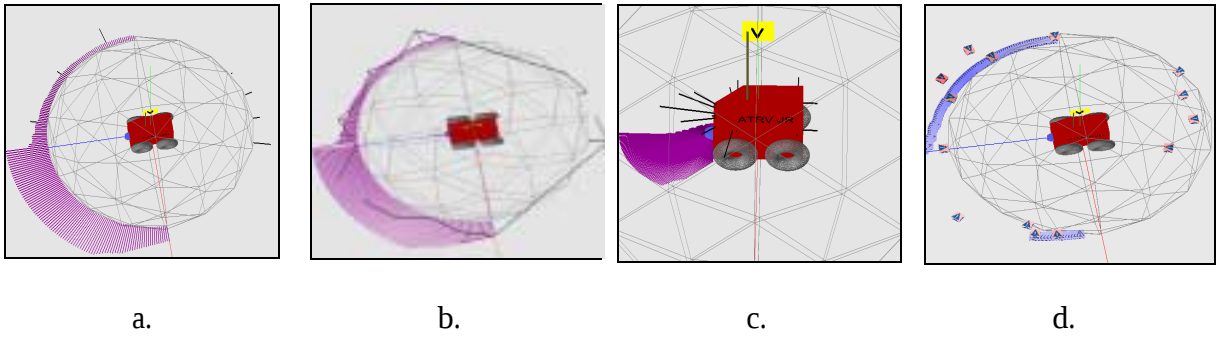


Figure 4.4. Sonar and Laser Ray Representation [Johnson, 2002]

In the graphical development of the SES, the camera data is placed on the tessellated dome nodes at the node closest to the point at which it was detected [Johnson, 2002]. In the initial implementation, well-known landmarks detected by the robot were represented as icons on SES. It is the intention is that actual camera images will be displayed on the SES dome at the location where they were located or as a virtual planetarium surrounding the robot. There is an assumption that it may also be more beneficial for the user to view sensory data from the robot's perspective. Due to this assumption, this work has added an egocentric view to the display option [Johnson, 2002]. Presently, there exist two views on the Sensory EgoSphere, the worldview and the egocentric view. Figure 4.5a shows object icons posted to the nodes of the SES. Figure 4.5b post panoramic images to the SES nodes. Figure 4.5c shows the panoramic images as in a planetarium view posted to the SES triangles. Finally, Figure 4.5c shows the same planetarium view from the robot's perspective.

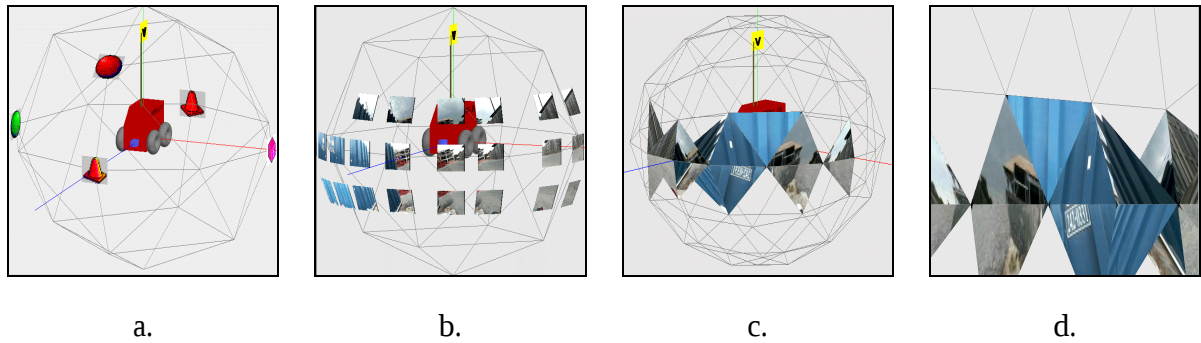


Figure 4.5. Camera Options and Views for the Sensory EgoSphere [Johnson, 2002]

### **Justification for the SES**

SES facilitates environmental event detection that simultaneously stimulate multiple sensors [Peters, 2001]. The SES provides a graphical display of the short-term memory database to the mobile robot user. This display is a compact illustration of the various sensory data surrounding the robot. The display provides the user with an egocentric perspective of the robot as it executes a task. The SES is a graphical representation of the SES database. The SES database is a searchable database that can be employed for extracting historical information regarding the robot's status during the mission or the last couple of minutes of activity. The SES can also provide a background behavior such as searching for an enemy during the task.

### **Applications of the Sensory EgoSphere**

#### **Supervisory Control**

In a supervisory control scheme, a person provides high level commands to a mobile robot. The robot then autonomously completes the commands. During, purely

autonomous activities the robot may be unable to autonomously complete a given task and may require supervisory intervention. In such instances, a more intuitive user-friendly display should assist the supervisor with resolving the situation [Kawamura, 2001a].

SES is a tool that a remote supervisor can employ to maintain awareness of the remote environmental conditions while sharing data between the robots in the field. Information regarding the current egocentric locations of known objects or landmarks within an environment can be very compactly coded with respect to a SES. Such a coding requires a label and a space-time location. This information can be transmitted to a supervisor or another robot, and either entity can develop its own SES. Over a low-bandwidth communication channel, the space-time position and label data can be transmitted to the supervisor in order to construct an iconic representation of a robot's environment. Broadband communications may enable a full immersion telepresence at the supervisor console [Peters et al., 2001].

Problems may still exist, even if the user has a presence in the robot environment. One such example is that at times disconnected numerical or graphical display of sensory data may not facilitate the user's pattern recognition skills [Kawamura, 2001a]. A solution to this dilemma is a directional, egocentric display, that is intended to maximize user's natural pattern recognition skills when combining sensory modalities. The proposal is that the addition of the SES to a graphical based human-robot interface should assist the supervisor's decision-making process while assisting the robot during difficult situations. SES may aid the supervisor when interpreting the robot's status. It is hypothesized that this system is an improvement over a mobile robot interface that only provides instantaneous feedback from unassociated sensors [Kawamura, 2001a].

The SES migrates information presentation from the sensing level to the perception level. The composition of the vision with other sensors on the SES surrounding the robot should provide clarity and ease interpretation. The SES should enable the user to better visualize the robot's present circumstances, while permitting the human supervisor to accurately ascertain the robot's present perception (sensory input) and employ such information while assisting the robot [Johnson, 2002].

The human-robot interface (HRI) is employed to provide the user with the robot's sensory information and status while providing a snapshot of the current environmental state [Kawamura, 2001a]. The HRI is implemented as a graphical user interface that contains the SES, a command prompt, a compass, an environment map, as well as sonar and laser displays. The hypothesis is that SES should enhance the supervisor's ability to understand the robot's circumstances and increase the supervisor's situational awareness [Johnson, 2002].

The IRL has a set of heterogeneous mobile robots that are coordinated by a human supervisor to accomplish tasks. The supervisor requires a robust HRI in order to manage the robot team. The current HRI research through direct sensor feedback has revealed a number of drawbacks. One disadvantage is that video communication requires high bandwidth, video storage, and high volume. Another disadvantage is that most mobile robots do not have 360-degree field of view. The user may encounter difficulty when combining diverse sensor information to accurately disseminate the robot's present surroundings and circumstances [Johnson, 2002].

The SES is a viable solution to some of these issues. The SES is considered to be a cognitive display because it represents the robot's short-term memory and displays the information graphically. During interaction with the world, the robot perceives the

environment and represents it in an egocentric manner. A secondary purpose of the SES is to provide a means to correct world perceptions by viewing the SES to detect misidentified or misplaced objects [Johnson, 2002]. The history feature of SES permits the user to replay the iconic representation of the sensory data. The amount of data retained in the history is limited by the hardware limitations. The history may assist the user when determining the robot's present state. The SES eliminates expensive video replay.

The composition of vision with the other sensors on the dome should provide clarity and interpretation ease to the user. The human supervisor communicates with the robot via the HRI that permits mission-level commands, provides an environmental map, laser display, sonar display, and the SES [Johnson, 2002]. It is proposed that the individual graphical representation of sensor agents does not provide the supervisor with a clear understanding of the robot's present state. Therefore, the SES is integrated into the interface. The consolidation of this data into one compact form should facilitate the user's access to a wide range of data. Real-time access to local sensor arrays, coupled with synthesized imagery from other databases, may provide the user with a virtual presence in an area from a remote location, thereby aiding the user with mission planning and other remote control tasks. The SES presents a compact display for various types of sensor arrays but is not considered to be sensory fusion. Sensory fusion develops a mechanism used to consolidate various modes of sensory data [Johnson, 2002].

The SES associates various sensing modalities and should greatly simplify the task of maneuvering a robot from trapped positions. The objects displayed on the SES also provide the supervisor with the ability to qualitatively command the robot rather than employing the traditional quantitative command mechanisms [Johnson, 2002].

## **Egocentric Navigation**

Dead-reckoning navigation using via points is notoriously difficult under any real-world circumstances due to the accumulation odometry errors [Kawamura et al., 2001b]. The robot must be able to react to actual surroundings and make adjustments to the route where necessary. The egocentric navigation approach employs via points to define via regions [Kawamura et al., 2001c, 2002]. A via region is a point on an a priori map with known landmarks surrounding it. The robot navigates reactively in via regions. The robot uses sensory data to navigate to each via region by aligning itself such that all landmarks surround it as given by the via region. [Kawamura et al., 2001b].

Experimental evidence suggests that as an animal enters a known locale, an egocentric description of the environment is developed in the hippocampus through interactions with long-term memory [Peters et al., 2001]. The long-term memory biases the hippocampus to respond to specifically anticipated sensory events.

An allocentric map (AMAP) is a map of the global environment that includes the relative locations of various landmarks [Peters et al., 2001]. The hippocampus response to sensory events depends on the interplay between egocentric and allocentric representations of the world. The egocentric map represents the objects' environment with respect to the animal at its current location. The AMAP is a Cartesian set representation containing directions that describe the larger scale relationships between locations [Peters et al., 2001].

The SES serves a similar purpose to the internal world representations. Suppose a robot is provided with an AMAP of the global environment that includes the relative locations of various landmarks. The robot can commence operation near a known location and the AMAP can be projected onto the SES to form a Landmark EgoSphere (LES) [Peters et al., 2001].

The LES can be referred to as a via point local map. By distributing the world map, each robot knows only its own environment. This representation minimizes the memory and computational requirements of each individual robot. When the operator defines a via point, a robot projects onto another EgoSphere the landmarks that it should be able to sense from the vicinity of the via points [Kawamura et al., 2001b].

The LES is a representation extracted from the long-term memory and is employed for localization using the current SES information [Kawamura et al., 2000a]. Figure 4.6 represents the robot's position on the world map and the generation of the Landmark EgoSphere extracted from the known landmarks found on the Sensory EgoSphere.

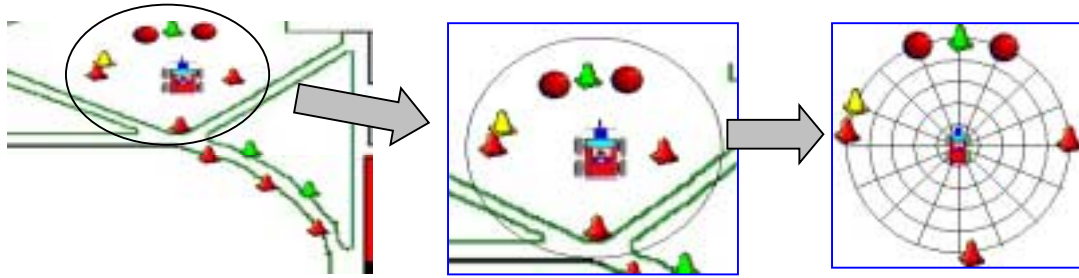


Figure 4.6. Landmark EgoSphere

The LES represents an egocentric map (EMAP) of the robot's location. The robot can determine its actual position within the environment given the angles from the robot's base frame to two or more of the expected landmarks localized on the SES. The robot centers itself by moving in the direction of the LES center. The robot continues to move until the objects on the great circle of the SES match the positions on the LES [Peters et al., 2001].

When given a via point by the operator, the robot projects onto another EgoSphere the landmarks that it should be able to sense from the vicinity of the via-point. This via-

point-local map is the LES. The robot reaches its next via-region by heading in the direction of a landmark visible from both the current location and the via-region while comparing the current contents of its SES to the LES of the target via-region. Tolerances are defined that permit the robot to identify the via-region even if the match between LES and SES is not exact [Kawamura et al., 2001c].

Global navigation is deliberative and local navigation is reactive. This division implicitly organizes the robot's memory into long-term and short-term components. The LTM is in part spatially organized to support global navigation. LTM represents the memory classification concept employed in egocentric navigation. Table 4.1 provides the functional classification of memory concepts.

Table 4.1. Functional Classification of Memory Concepts [Kawamura et al., 2001b]

|                | Working Memory                   | Short-term memory                | Long-term memory         |
|----------------|----------------------------------|----------------------------------|--------------------------|
| Storage        | Robo-centric topological regions | Robo-centric topological regions | Global layout            |
| Representation | Via-regions                      | Sensory EgoSphere (SES)          | Landmark EgoSphere (LES) |
| Implementation | Global                           | Local                            | Global                   |
| Response       | Deliberative                     | Mostly Reactive                  | Deliberative             |
| Persistence    | One task                         | A limited number of tasks        | Several tasks            |

At any specific environmental location, the sensory horizon defines the region that the robot can sense. Only the objects within the region have the possibility of being sensed and stored on the SES. During navigation, the robot periodically updates its STM. The updates result in the creation of SES structure instances at discrete locations. Each SES

instance represents an environmental snapshot at a specific space-time location. These discrete structures form a SES representation chain that defines a topological map. At navigation completion, a series of SES regions are stored in the short-term memory (See Figure 4.7).

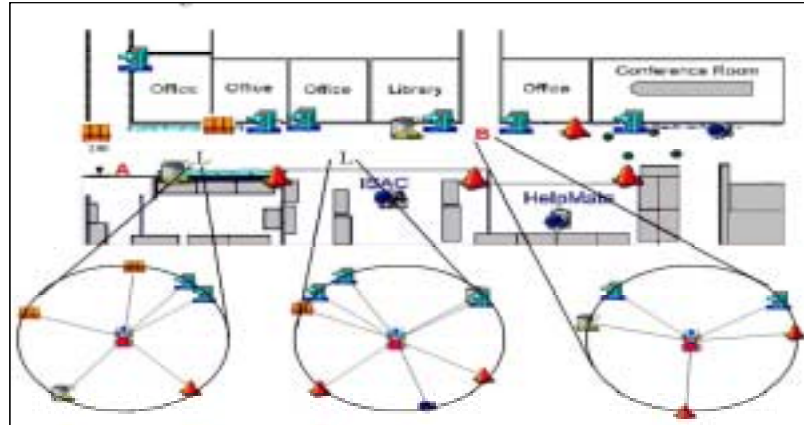


Figure 4.7. SES representation chain [Kawamura et al., 2002b]

In addition to the long term and the short-term memory, the robot also maintains a working memory. The working memory contains the descriptors of key locations or via regions that indicate robot navigational points. Navigation depends upon a sketch that is qualitative and is neither metrically precise nor accurate but is descriptive enough so that humans can follow certain landmarks and the target location. Sketches do not pinpoint the target but indicate a close proximity to the target. The sketch's inaccuracies are compensated for by perception of the actual scene and the user's reasoning capabilities. Humans possess rich sensing capabilities and high intelligence, therefore humans are better able to use the sketches than the robots. This egocentric approach avoids the need for distance information while navigating [Kawamura et al., 2001c; Kawamura et al., 2002b].

Researchers studying traditional robot navigation seek high precision navigation that results in sensing and actuation brittleness or fragility. Egocentric navigation employs regions in navigation and localization. Reaching a certain region is an imprecise action; performing an action within this region might require more precise localization of the robot [Kawamura et al., 2001c].

Perception represents the robot's ability to develop an abstraction based on sensory readings and the robot's ability to recognize a limited number of objects. This perception may occur by employing sonar, laser, and vision information. Object descriptors are necessary and an object library is implemented. The robot stores certain object descriptors to the library and retrieves them when necessary [Kawamura et al., 2001c].

Humans tend to rely upon angular information rather than distance information while learning places and localizing themselves. Human behavior is consistent with the behavior in egocentric navigation. An egocentric representation is used for describing the current robot situation and describes potential paths and target readings. A region is described by the landmarks that are visible or are expected to be visible from within that region [Kawamura et al., 2001c].

Angular representation is the basic component of the entire egocentric navigation architecture. The representation is referred to as SES or LES based upon the information source. If the representation created is based upon perception, it is termed a SES. If the representation is extracted from the robot's a priori map or is provided by a user or another robot as a target location, it is termed a LES. The SES and LES provide basic range-free egocentric navigation since they are based solely on angular information rather than on metric maps and distance information [Kawamura et al., 2002b].

In conclusion, this chapter presents the foundations for the SES. The sensor egosphere was first defined by Albus in 1991. In 2001, the SES was defined by the IRL as a discrete tessellated dome that represents the STM of a robot. This chapter presents SES as a short term memory, the graphical representation, and applications for the SES. This work migrates the SES in theory to an agent that is realized in a graphical user interface.

## **CHAPTER V**

### **GRAPHICAL BASED HUMAN ROBOT INTERFACE**

This chapter introduces the agent-based Intelligent Machine Architecture (IMA) [Pack, 1998], a robot control architecture software used in the Intelligent Robotics Laboratory that allows the concurrent execution of software agents on distributed machines while facilitating extensive inter-agent communication. The chapter then describes the development of a graphical based human-robot interface under IMA. The key components of the HRI include the SES agent, the map agent, the sonar agent, the laser agent, and the camera agent. These components were developed as part this research.

#### **Introduction**

Intelligent Machine Architecture (IMA) is a multi-agent robot control architecture [Pack, 1998]. IMA provides a means for developing software agents that communicate in a distributed computing environment. The IMA agents consist of components, atomic agents, and compound agents. The components are Microsoft DCOM objects that enable communications between agents and are building blocks of the atomic agents. There are five types of atomic agents: hardware/resource, behavior/skill, environment, sequencer, and multi-type [Pack, 1998]. The hardware/resource agent interfaces to sensor or actuator hardware. The behavior/skill agent contains basic robot behaviors or skills. The environment agent provides an abstraction of the robot's environment. The sequencer agent performs a sequence of operations, often on one or more atomic agents. The multi-type agent combines the functionality of at least two or more of the previous agent types.

Finally, the compound agent is an interacting group of atomic agents that are coordinated or sequenced by one or more sequencer agents. The SES agent primarily communicates with the sensor and actuator agents that provide abstractions of sensors and actuators while incorporating basic processing and control algorithms [Kawamura, et al., 2001a]. Pack [Pack, 1999] provides a more in depth study of IMA. Table 5.1 provides a listing of the behaviors currently available on the ATRV-JR robot. Some of these behaviors were modified or updated for use in the evaluation studies.

Table 5.1. Agent Structure of the ATRV-JR robot [Thongchai, 2001]

| Agent       | Method                                                                                                                                                                   | Comments                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Modifications                                               | Used<br>(Y/N) |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|---------------|
| Avoid Enemy | Laser data is employed to determine angle, and distance to an object. The force values are determined by calculating linear and angular velocity to turn the robot away. | The enemy is determined by object size and is detected by the laser. The user provides the range of object size to denote as the enemy. Includes Runaway, Collide, Avoid, Avoid Static obstacle behaviors. Runaway converts a total enemy force into linear and angular velocity. Collide immediately stops the robot if within a certain range. Avoid takes the repulsive force from the sonar and combines with the desired heading to produce a new heading. | The enemy will be a certain object distinguishable by color | Y             |

Table 5.1, continued

|                |                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |      |   |
|----------------|--------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|---|
| Avoid Obstacle | Potential Fields                                       | The robot is attracted to its goal and repulsed by obstacles in a virtual potential field. The sum of the sonar readings is used to create a resultant force field. This is summed with the vector from the robot to the target to get a net vector force. Includes Runaway, Collide, Avoid, Avoid Static obstacle behaviors. Runaway converts a total enemy force into linear and angular velocity. Collide immediately stops the robot if within a certain range. | None | Y |
| Emergency      | Range sensor and minimum closest distance.             | If the distance between an object and the robot is less than a minimum distance then the robot will completely stop. The minimum distance decreases as the speed of the robot increases. If the robot gets close to the minimum distance in the front then it checks for an obstacle behind it and moves back. If it is already within the threshold, then it will stop                                                                                             |      | Y |
| Follow         | Follow Wall<br>Follow Corridor<br>Follow Moving Object | These behaviors are based upon the same potential fields as the avoid obstacle.                                                                                                                                                                                                                                                                                                                                                                                     |      | N |
| Motor Control  | Turn Left<br>Turn Right<br>Forward<br>Backward         | These are the basic commands to drive the motors on the robot's base.                                                                                                                                                                                                                                                                                                                                                                                               |      | Y |

Table 5.1, continued

|              |                                                         |                                                                                                                                                                                                                          |  |   |
|--------------|---------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|---|
| Move to Goal | GPS, fuzzy control, dead reckoning                      | The robot can autonomously move to a predefined goal using a series of way points, implemented with GPS and fuzzy control. Includes move to point, move to GPS Point, Move to Box, Move to Cone, Move to Ball behaviors. |  | Y |
| Perceptual   | Detect Obstacle<br>Detect Target                        | Typically done using color                                                                                                                                                                                               |  | Y |
| Wander       | Generates a new heading for the robot every 10 seconds. | Uses some type of random number generator to change the robots angular and linear velocity                                                                                                                               |  | N |

The ATRV-JR sensor suite has an odometer that provides the robot's position (x, y) and the heading angle relative to the robot's initial position. It also has sonar that transmits ultrasonic signals and measures the time of flight of the returning signal from obstacles. The laser sensor is mounted on the front of the robot and scans the environment by sending out laser from 0 to 180 degrees and measuring the returned signal that indicates the distance to the object. The pan-tilt-zoom Sony camera system integrates a high-speed range of  $-100$  to  $110$  degrees. The second camera is a USB camera used to pan the rear of the robot. The robot also contains a compass, GPS, and DGPS sensors. The ATRV-JR's current interface can be divided into the hardware interface and the IMA agent. The resource agents are the base, odometry, sonar, laser, GPS, DGPS, compass, and power. The behavior agents are avoid-obstacle, avoid-enemy, and the others are listed in Table 5.1. Figure 5.1 displays the sensor suite for the ATRV-JR robot. The left side of Figure 5.1 shows the dimensions of the sixteen sonar around the perimeter of the robot as well as the laser in the front of the robot.

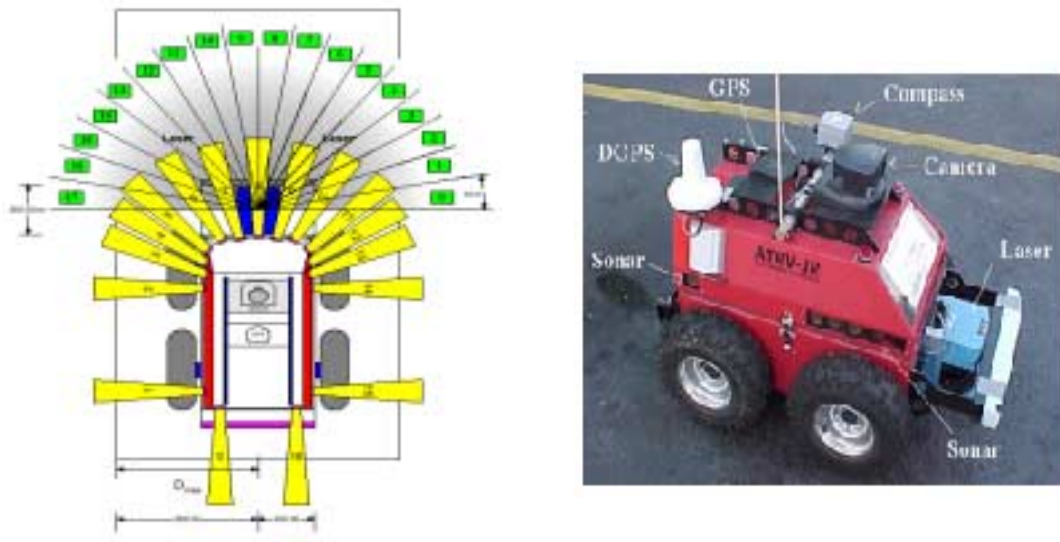


Figure 5.1. Sensor Suite on the ATRV-JR robot

The Multi-Agent Based Robot Control Architecture includes a SES, LES, Self (Robot) Agent, Commander Interface Agent, EgoSphere Manager, Database Associative Memory (DBAM), and DBAM Manager. The two compound agents are the Commander Interface Agent and the Robot Agent. These two agents represent the two cognitive agents for the agent-based human-robot interface. The Robot Agent is used to monitor how an agent achieves a task and how each sensor is working and then share that information with the human operator. The Robot Agent also receives commands from the human and takes appropriate action. Figure 5.2 represents the Multi-Agent Based Robot Control Architecture for the ATRV-JR robot.

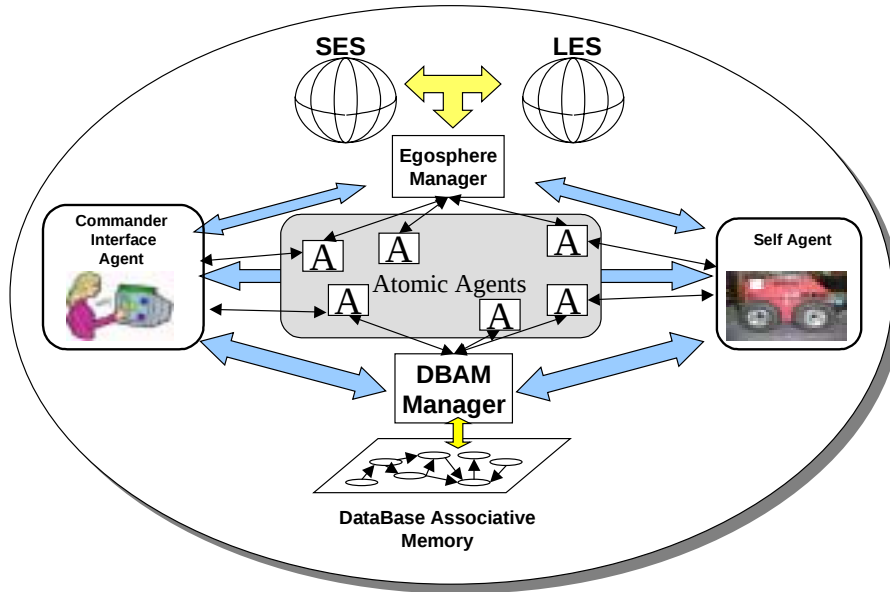


Figure 5.2. Multi-Agent-Based Robot Control Architecture [Kawamura et al., 2001a]

The Landmark EgoSphere (LES) in Figure 5.2 is the data structure that contains the topological map of the known landmarks. The LES provides the user with landmarks the robot either acquired through its sensor suite or were specified by the user. The Data Base Associative Memory (DBAM) provides the long-term memory for the robot. It is composed of data to enable the robot to recall sequences of actions based on the state of its internal and external environments. The records in the database are relational and are associated by weighted links. The records contain action instructions, object descriptions as well as other information needed by the robot. This mechanism incorporates the robot's domain knowledge, sensory input, and the physical status be modified by the Spreading Activation (SAN) through reinforcement learning. This architecture was extended to include the human-robot interface through the graphical user interface, off line mission planning, and the user command post [Nilas, 2003]. Figure 5.3 shows the integrated Agent-based Human-Robot Interface and control architecture.

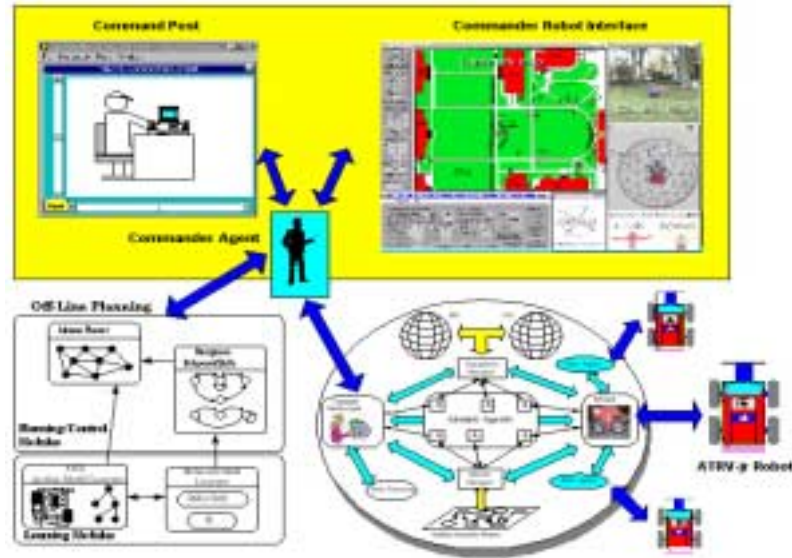


Figure 5.3: Integrated Agent-based Human-Robot Interface and Control Architecture

### **Design Overview (Enhancing a HRI)**

#### **SES Agent**

The SES agent designed in IMA, which included the graphical geodesic dome, was added to the original human-robot interface (see Figure 6.1). This agent is not only a graphical display of the robot's sensory data but is a simplified representation of the robot's short-term memory. This agent communicates with the robot's other sensory agents as well as the other HRI agents. The other interface agents include the sonar, laser, compass, and camera.

This chapter presents the background information for the agent-based graphical user interface, under IMA. In order to demonstrate the concept of enhancing an HRI using an SES, a very basic user interface was implemented using this architecture. The basis interface includes the following agents: SES, laser, sonar, camera, map, and command. The second

phase of this research involves the evaluations of the enhanced system. The experiments and test bed are presented in Chapter VI.

## CHAPTER VI

### EXPERIMENTAL DESIGN

The research proposal includes the development of the graphical based Human-Robot interface and the SES agent presented in Chapter V. This chapter focuses on the experimental design for the evaluation of the enhanced interface. This chapter presents the research questions, the goal, and the hypotheses associated with this work. The experimental test bed, tasks, evaluation methods are also included.

#### Introduction

It is proposed that the addition of the SES to an agent-based Human-Robot interface will enhance the interface usability. It is also hypothesized that the addition of the SES will increase the participants' situational awareness and reduce workload for supervisory control of mobile robots.

#### **Research Questions**

The general research question can be stated as follows, *Can the addition of the SES to a HRI improve the participants' ability to supervise mobile robots?*. More specifically, *Can the addition of the SES to an HRI,*

1. *Improve the GUT's usability?*
2. *Improve the participants' situation awareness?*
3. *Decrease the participants' workload?*
4. *Decrease task execution time?*

## **Research Goal**

The research goal is to develop a more effective and efficient graphical based human-robot interface based upon an agent-based framework with the addition of a Sensory EgoSphere for supervisory control of mobile robots.

## **Research Hypotheses**

The aforementioned research questions and the goal can be summarized in two hypotheses. They are,

1. The SES decreases participant mental workload with the addition of a more intuitive display of sensory data.
2. The SES increases participant situational awareness of the robot status and the task/mission status.

In order to verify the aforementioned hypotheses a set of experiments were designed that require the participant to accomplish a given task using the HRI. The experiments included participants with varying levels of education, experience with graphical user interfaces as well as mobile robots. The data collection included the participants' perceived workload, perceived interface usability, and task execution time. Some of the questions this study was designed to answer were:

- Does the enhanced HRI decrease the task execution time?
- Does the enhanced HRI decrease errors in participant responses?
- Does the enhanced HRI increase the participants' situational awareness?
- Does the HRI increase the interface usability?
- What specific components of the HRI did the participant access most frequently during task execution?

## **Test Procedure**

### **Summary of Procedure**

This following is a brief overview of what the system evaluation entailed. The first stage included an orientation followed by a training session. During the training session, the participants became familiar with the interface components, including the sonar, laser, camera, and SES. In the training task, the participants used the interface to find the robot. For the second set of tasks, the participants performed simple navigation tasks by providing high level commands to the robot via the interface. During the task execution, quantitative data was collected in the form of videotaping, automatic data recording, and a spatial reasoning test. Additionally, subjective data was collected via questionnaires and cooperative evaluation.

### **Place and Time**

All evaluations took place inside Vanderbilt University's Featheringill Hall. The participants used the human-robot interface located in the evaluation room. The mobile robot was located in the hallway out of the participants' view. The time evaluations occurred during a typical workday.

### **Participants**

The study participants included eleven novice and sixteen experienced participants. Novice participants are defined as participants with little to no experience with video games, mobile robots, user interfaces, and computers. The experienced participants had some familiarity with robots, computers, and teleoperation of remote systems. The study included

seventeen female and ten male participants. The justification for using novice and expert participants was to confirm that novice participants were able to effectively use the interface to extract vital information about a robot, as well as send commands to the robot. Cohen proposed the following methodology for selecting participants [Cohen, 1995]:

"Participants should include a range of skill levels in order to control for the possibility that high performance is due to easy problems. The inclusion of the novice participants represents control group of problem solvers who can solve easy problems but not difficult ones. The addition of both levels of experience also sets a higher standard by which the performance is measured."

### **Equipment and Materials**

The test site included one desktop personal computer, one laptop computer, a television monitor, a keyboard, a mouse, one table and two comfortable chairs. There were also two video cameras for recording the participant and computer during task execution. The remote site included one mobile robot, the ATRV-JR. The materials required were two consent forms, the training materials, the pre-/post-experimental questionnaires, post-task questionnaires, spatial reasoning test, subjective workload assessment, and environment maps.

### **Experimental Design**

The participants were categorized by computer experience, knowledge, skill, spatial reasoning ability, gender, and age range. The spatial reasoning was determined by the administration of a spatial rotation test. The spatial rotation test determines the participants' spatial relationship capabilities. Since the participants operated a mobile robot via an

interface from a remote location, the ability to visualize three-dimensional relationships is important. On the spatial reasoning test, scores between zero and four are considered low, scores between five and eleven are considered medium, and scores between twelve and twenty are considered high. The other characteristics determined from a pre-experiment questionnaire, included familiarity with computers, video games, mobile robots, and graphical user interfaces. At the task completion, the participants completed a post-task questionnaire to evaluate their perceived workload. At the conclusion of the evaluation, the participants also completed a post-experiment questionnaire.

The independent variable in an experiment is the feature that is deliberately varied by an experiment and the dependent variable is what is measured in the experiment. One dependent variable for this study is the task execution time. The perceived participants' workload was also a dependent variable. The independent variable was the changes in the components of the user interface screen, such as the addition of the SES. It was assumed that the addition of the SES would enable the participant to more quickly assess robot circumstances, therefore reducing the task execution time. The additional information provided by the SES should decrease the amount of effort the participant must exert to accomplish a task. Consequently, the task execution time and participants' mental workload are dependent upon the components of the interface.

### **Data Collection**

The data collection included videotaping of the participants as the tasks were executed. Upon task completion, the participant viewed the videotape and performed a cooperative evaluation of their task performance. Examples of questions asked were: "where do you think the robot is on the environmental map?", "what object did the robot

just pass?”, “how close do you think the robot is to the goal?”, “what do you think the sonar lines represent?”, “Can you correlate data from individual sensors to data on the SES graphic?” With respect to the quantitative data collection, the participants were evaluated on how accurately they can determine mission status and the robot’s progress based upon the interface components.

A second method of quantitative data collection included automatically recording the mission start time as well as the mission completion time along with the number of and type of participant mouse clicks and command errors. Each task was executed twice, once with the original interface and once with the enhanced interface.

The participants also completed a subjective workload assessment to evaluate the participants’ workload during the task execution. Subjective measures require the participant to rate their perceived workload, typically their feeling of exertion and effort during task execution.

The participants also completed several questionnaires, including a pre-experiment, post-experiment, and post-task. In all, there was one pre-experiment, one post-experiment, and four post-task questionnaires.

The final evaluation data collected was a comparison of the participants’ before and after performance after completing a task the second time, once with the original interface and once with the enhanced interface. Appendix G includes a full description of tasks A, B, E, and F. Task A and E did not include the SES, while Task B and F did.

Data collection topics included usability, situational awareness, mental workload, and interface components. There were two people and two video cameras for the usability testing; one set for the participant and one set for the mobile robot.

### **Schedule**

Each participant was required to commit to completing two 75 minutes sessions. The schedule for the first day of data collection is shown in Table 6.1. Table 6.2 represents the schedule for the second day.

Table 6.1. Session 1 Schedule

|                                |            |
|--------------------------------|------------|
| Orientation                    | 5 minutes  |
| Training                       | 15 minutes |
| Pre-Experimental Questionnaire | 5 minutes  |
| Tasks                          | 40 minutes |
| Video Review                   | 5 minutes  |
| Post-Tasks Questionnaire       | 5 minutes  |
| TOTAL                          | 75 minutes |

Table 6.2. Session 2 Schedule

|                               |            |
|-------------------------------|------------|
| Welcome/ Training Review      | 10 minutes |
| Tasks                         | 40 minutes |
| Video Review                  | 10 minutes |
| Post-Tasks Questionnaire      | 5 minutes  |
| Post-Experiment Questionnaire | 10 minutes |
| TOTAL                         | 75 minutes |

The data collection process took place over approximately one month from July 30, 2002 to August 23, 2002. Depending on the availability of participants, two to four individuals completed the study per day.

## **Design Overview**

The evaluations for this work were task-dependent in order to determine the advantages and disadvantages of the SES for various tasks. The experiment employed scenarios during which participants were asked to accomplish certain tasks. The participants evaluated the HRI based upon this task. During the task execution, the participant and the robot share autonomy. Each participant was provided a 15-minute training session in which the interface components were explained and the participant had an opportunity to view robot navigation via the interface. The order in which each participant completed the tasks was randomized. The participants performed the training task twice and the evaluation task twice, once with the original interface and once with the enhanced interface. The prototype for both of these interfaces shown in Figure 6.1.

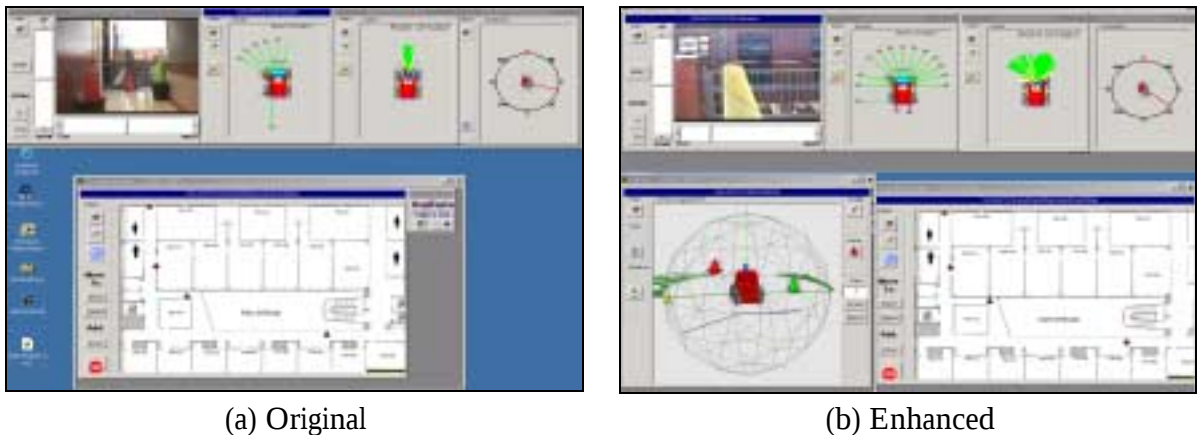


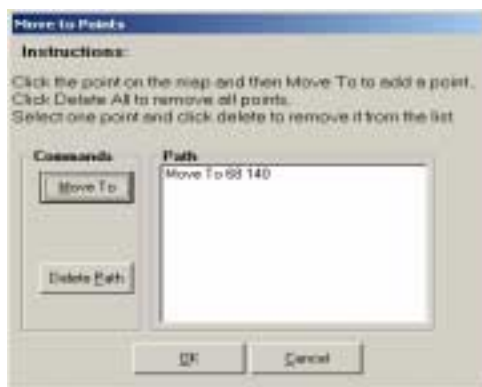
Figure 6.1. Prototype of Human-Robot Interfaces

Scooter, the ATRV-JR robot, was equipped with two cameras that provide pan-tilt-zoom capability. In order to generate the SES, the participants must issue a command for the cameras to pan the environment and update the graphic. The sonar and laser data

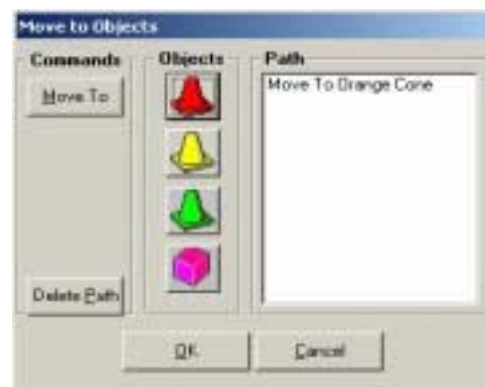
update continuously, but the images update only upon command. The robot stopped all movement to collect the imagery data and alleviate camera vibration.

The low-level commands to the ATRV-JR's base include: drive go straight, turn left, turn right, and stop. These commands can be combined with the robot's sensory data to develop basic behaviors. Scooter has several available behaviors including find object, move to point, move to object, and avoid obstacle. The obstacle avoidance employs potential fields. During training, the participants learned the move to point, move to object, and avoid obstacle behaviors. During the interface evaluation, the participants used the move to object and move to point commands.

The participants input navigation commands using point and click interaction on an environment map. The move to point method involved using the mouse to click on the via points and then commanding the robot to move to that point. The participants selected icons on the move to point screen to command the robot to move to an object. Figure 6.2 exhibits prototypes of the two command input options.



(a) Move to Point



(b) Move to Object

Figure 6.2. Navigation Command Input Options

## Situational Awareness

Situational awareness is the knowledge of what is going on around the human operator or the robot. In this work, the analysis of SA is restricted to the awareness of the human operator. There are three levels of situational awareness: perception, comprehension, and prediction. Perception is the use of sensors in order to determine the surroundings. Comprehension is using perception to become aware of the robot situation. Prediction is a higher level of comprehension that uses present circumstances to determine what the robot will do next [Endsley, 1989]. It is proposed that the Sensory EgoSphere would move the participant from the perception level to the comprehension level. Figure 6.3 demonstrates the three levels of situation awareness. Situation Awareness was measured by comparing the participants sub-task and task scores with both interfaces.

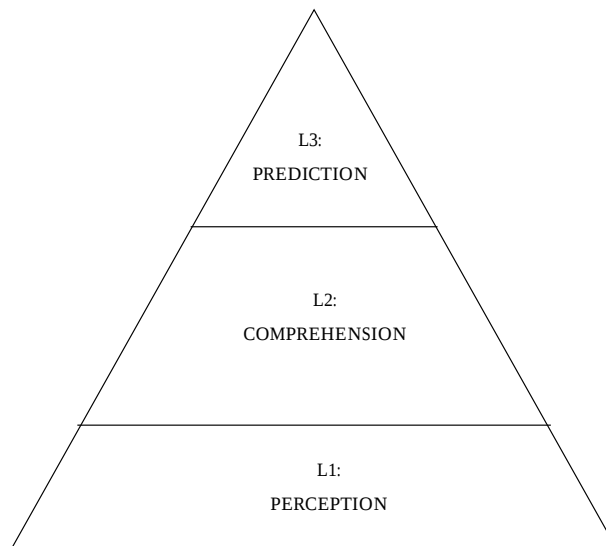


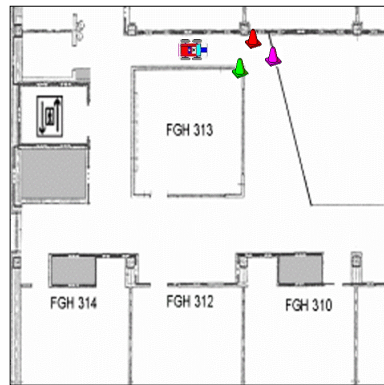
Figure 6.3. Three Levels of Situation Awareness [Endsley, 1989]

### **Task Scenarios**

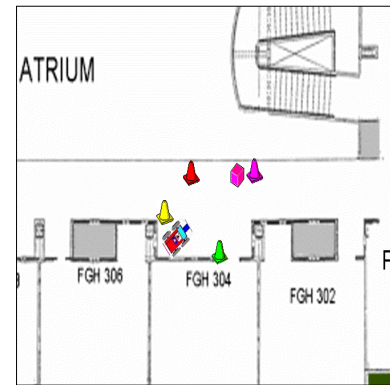
The participants were required to perform four tasks using the human-robot interface. The first two tasks were training tasks. In the training tasks, the participant searched for the robot by using all of the interface displays. In the second set of tasks, the participant performed the evaluation of the interface while teleoperating the robot. In these tasks, the robot was sent a navigation command and as the robot traverses the path, data was recorded. Each task was completed twice, once with the original interface and once with the enhanced interface (see Figure 6.1). During all tasks, the participants were given the option of changing the SES as well as other sensory display views. All tasks were used to test the hypotheses concerning situational awareness and mental workload.

#### **Task One: Find the Robot (Training)**

In task one, the participants became familiar with the interface by using it to find the robot. The participant was told that the robot was located on the third floor of the engineering building. The participant employed all of the display screens to locate the robot. The participants panned the environment to locate significant landmarks around the robot and recorded the robot's location on a printout of the map. The participants also recorded all objects found around the robot. Finally, the participants described how to drive the robot back to the home position. Figure 6.4 provides the environment map layout for the training task while Table 6.3 is the task allocation.



(a) no SES



(b) with SES

Figure 6.4. Scenario One

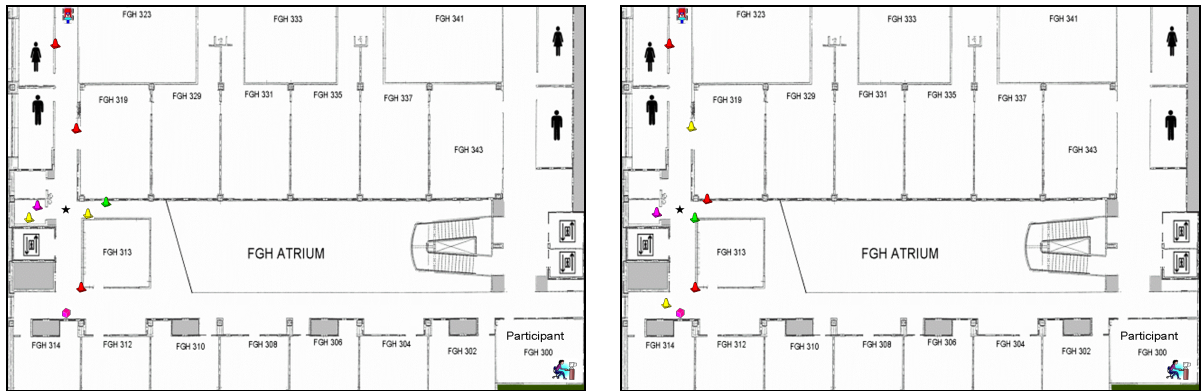
Table 6.3. Scenario One task allocation

| Step | Task                                   | Participant | Robot |
|------|----------------------------------------|-------------|-------|
| 1    | Use the camera to scan the environment | X           |       |
| 2    | Issue command to scan the SES          | X           |       |
| 3    | Update the SES or camera view          |             | X     |
| 4    | Record location of all objects         | X           |       |

### Task Two: Drive the Robot (Evaluation)

In task two, the participants supervised a mobile robot as it executed a high-level command. The participants provided via points, objects, and a goal point on a path for the mobile robot to execute. In order to reach the goal, the robot passed through several via points and avoided obstacles. The participants were provided a printout of the floor plan.

The participants recorded objects passed while the robot moved to various points. The participants used the camera to view the environment and locate objects at each via point. For the enhanced HRI, the participants also issued an SES scan request. There were certain tasks allocated to the participants and the robot, these are delineated in Table 6.4. Figure 6.5 represents the graphic of the two scenarios for the evaluation task. The difference in the two scenarios is the location and color of the landmark objects along the path. Appendix G provides the detailed instructions for the evaluation tasks.



(a) no SES

(b) with SES

Figure 6.5. Scenario Two

Table 6.4. Scenario 2 task allocation

| Step | Task                                                     | Participant | Robot |
|------|----------------------------------------------------------|-------------|-------|
| 1    | Use the Move to Point command to move the robot          | X           |       |
| 2    | Find objects and move to objects on path                 |             | X     |
| 3    | Signal the participant upon arrival to objects or points |             | X     |
| 4    | Use the camera to scan the environment                   | X           |       |
| 5    | Issue the scan command                                   | X           |       |
| 6    | Scan the environment to update the SES or camera view    |             | X     |
| 7    | Record all objects found on the landmark map             | X           |       |

## **CHAPTER VII**

### **DATA ANALYSIS AND RESULTS**

This chapter summarizes the results of the human factors evaluation. As previously stated, each participant performed four tasks, two of which were training tasks. Training Tasks A and B involved determining the robot's position in a specified area using the interface displays. Task B included the SES display while Task A did not. Tasks E and F entailed driving the robot through an obstacle course while documenting all significant objects passed. Task F included the SES display and Task E did not. This chapter discusses the statistical analysis that was conducted as well as the results.

#### **Participant Demographics**

Of the 27 participants who completed the interface evaluation, there were seventeen female and ten males. The spatial reasoning test showed that seven participants had low spatial reasoning, six participants had high spatial reasoning, and fourteen participants had average spatial reasoning. The average participant age was 30 years old. There were nine participants between the ages of 19 and 20, thirteen participants between 20 and 40, three between 40 and 60 and two participants over 60 years. There were twelve undergraduate students, one graduate student, and one high school student. Of the students, eight had non-technical majors such as business and psychology. There were three elementary school teachers and one biology professor. Included in the remaining occupations were two engineers, two accountants, one lawyer, one nurse, one physical trainer, one social worker, and one building manager. Due to several issues with time conflicts and failures, many

participants were unable to complete all tasks. There were seven participants who completed all four tasks, six participants completed three tasks, five participants completed two tasks, and nine participants completed one task. The failures will be discussed in more detail in the next section.

## **System Failures**

### **Automatic Data Recording**

All data was recorded via an electronic questionnaire that recorded responses and exported the responses into an Excel spreadsheet. In certain instances, the application failed to open or write the data, therefore some of the data was lost. Overall, there were seven automatic data recording failures, four sets of post-experimental and three sets of post-task recording failures.

### **Hardware**

There were four hardware problems with the robot during the course of the interface evaluations. The first issue involved the wireless network security settings changing during one participants' evaluation, the result was a communication failure between the interface and the robot. This failure took an entire day to repair. The second issue occurred during the robot movement, the wireless card disconnected from the robot and required one hour to diagnose and repair. Additionally, the camera cable had a bad connector and therefore the camera commands were being intermittently received. This problem took one day to diagnose and repair. The final dilemma was that although the robot had a front Sony PTZ camera and a rear USB camera, the rear camera did not update the camera view. This failure

meant that the participant could view only 90 degrees to the left and right and that the SES only detected objects in the front hemisphere.

### **Software**

The main software failure involved the camera. Due to the hardware problem previously discussed, at times the participants could see the robot's view but could not operate the camera. The failures usually involved either a delay or complete failure in the execution of pan, tilt, and zoom commands. The second most prevalent failure was with the compass. Due to the magnetic field in the engineering building, at times the compass display would have an error in the heading of up to 90 degrees. The final error was a very typical one in robotics research. Because of the robot's odometry error, the robot's position on the map began to diverge greatly from the robot's actual position. In light of all three of these software failures, it was necessary for the participants to compensate for these issues when making decisions about the robot's position and surroundings.

### **Statistics**

Twenty-seven participants completed the defined tasks employing the human-robot interface during the first session. Several participants were unable to return for a second session while others encountered software or hardware failures during their visit. Due to these challenges, the sample size for the data analysis was reduced to ten participants. These ten participants had completed Tasks E and F without any major system faults. Of these ten participants, six also completed both Tasks A and B, while three completed A and one completed B. There were actually twelve participants who completed both Tasks E and F but two had failures recording the data from their post-task and post-experiment

questionnaires. Due to these failures, these two participants were eliminated from the analysis set.

In the ten participants, there were five males and five females. In this group, there were two participants with low spatial reasoning, four with average reasoning, and four with high reasoning. With respect to the ages, there were five participants between 18 and 20, four between 20 and 30 and one participant over 30. In this group, there was one graduate student, five undergraduate students, one high school student, one nurse, one building manager, and one accountant. Of these ten participants, seven had non-technical majors or occupations.

Due to the small sample size, neither normal distributions nor equal population variances were appropriate for the data analysis. Therefore, all non-parametric tests were performed. There are a number of issues associated with non-parametric test. Such issues include reduced sensitivity, the use of less information, and less efficiency than their parametric counterparts. Hypothesis testing was employed to test theories regarding the addition of the SES to the human-robot interface. Some of the hypotheses to be tested were that the SES would reduce workload, decrease task time, and increase situational awareness. Correlation testing was employed to determine if there was a direct relationship between task performance and other participant qualities such as mouse clicks, spatial reasoning or overall satisfaction. All analyses were evaluated using either Microsoft Excel macros or SPSS and many were verified using both techniques.

### **Hypothesis Testing**

“A statistical hypothesis is a conjecture about a population parameter. This conjecture may or may not be true [Bluman, 1998]”. Hypothesis testing is performed by

assuming that there is a null hypothesis for a sample data set. The null hypothesis is the assumption that all of the sample data comes from the same population. The null hypothesis is usually denoted by  $H_0$ . There are three types of tests, two-tailed, right-tailed, and left-tailed. The two-tailed test indicates that the null hypothesis should be rejected if the test value for the population is in the critical region. The left and right-tailed test is more specific as to whether the test value is to the left or right of the critical region. All of the statistical analyses for this evaluation employed the two-tailed test. The level of significance, p-value, used for all the tests in this study was 0.05. This level of significance means that there is only a 5% chance that samples with these characteristics came from the same population. Therefore when the p-value is less than or equal to 0.05, it can be assumed that the null hypothesis is rejected.

### ***Kruskal-Wallis Rank Test***

The Kruskal-Wallis Rank test is a non-parametric test that is sometimes referred to as the H test. This test can be employed to compare three or more means. The H test distribution can be approximated by the chi-square distribution with k-1 degrees of freedom. This test uses the data ranking to determine the validity of the null hypothesis. All the data is considered as a group and is ranked. The H formula is then used to distinguish the ranks. The H formula is an approximation of the variance of the ranks. If the samples are from different populations, the sum of the ranks will be different and there will be a large H value. If the samples are from the same population, the H value will be small and the null hypothesis will be rejected. The H formula is given by,

$$H = \frac{12}{N(N+1)} \left( \frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \dots + \frac{R_k^2}{n_k} \right) - 3(N+1)$$

where  $R_k$  is the sum of the ranks of sample  $k$ ,

$n_k$  is the size of sample  $k$ , and

$$N = n_1 + n_2 + \dots + n_k$$

$K$  = number of samples that will be compared.

### ***Wilcoxon Rank Sum Test***

The Wilcoxon rank sum test is employed for independent samples while the Wilcoxon signed-rank test is employed for dependent samples. The parametric counterparts of these two tests are the Z-test and the matched-pair T-test. The only assumption for the Wilcoxon tests is that the population of differences is symmetric. In this test, all the data is combined and ranked. When all the ranks for each sample are summed, if the sums are approximately equal, then the null hypothesis will not be rejected. The formula for the Wilcoxon Rank sum test for independent samples is given here.

$$z = \frac{R - \mu_R}{\sigma_R}$$

where,

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} \quad \sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

$R$  = sum of the ranks for the sample size ( $n_1$ ),

$n_1$  = smaller of the sample sizes, and

$n_2$  = larger of the sample sizes.

### ***Wilcoxon Signed-Rank Test***

The Wilcoxon Signed-Rank Test is employed when population samples are dependent. This test can be used in place of the T-test for dependent samples. In this

procedure, the samples are matched and the difference between the samples is ranked. Each rank is assigned the sign of the difference. The sum is found for the positive ranks and the negative ranks. The sum is then compared to the test value,  $w_s$ . If the value of the sum is less than the test value, then the null hypothesis is rejected. The formula for the Wilcoxon Signed-Rank Test is,

$$z = \frac{w_s - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

where  $n$  = number of pairs where the difference is not zero,

$w_s$  = smaller sum in absolute value of the signed ranks.

### ***Friedman Test***

The Friedman test is a non-parametric two-way analysis of variance statistic. This statistic compares the means for two or more related samples. The Friedman test ranks the values of each matched set (row) and then sums the rank of each group (column). If the sums of the ranks are very different, then the p-value will be small. This matched test is used to control for experimental variability between participants. This test is a function of the sums of the squares of the deviations between the rank sums. The test statistic is given by,

$$S = \sum_{j=1}^n \left[ R_j - \frac{k(n+1)}{2} \right]^2$$

where  $k$  = the number of related samples (rows),

$n$  = the number of treatments with each of  $k$  blocks (columns),

$R_j$  = the sum of the ranks for each treatment, and

j = the jth treatment .

### Correlation

Correlation is employed to determine whether a relationship between two variables exists. Regression is used to determine whether there is a positive or negative relationship. If there is a relationship, the correlation coefficient determines the strength of the relationship between the two variables. The symbol for the sample correlation coefficient is r and for the population correlation coefficient, it is  $\rho$ . The range of the correlation coefficient is +1 for a strong positive relationship to – 1 for a strong negative relationship.

Scatter plots are usually employed to determine the relationship between the independent and dependent variables. A regression line is drawn through the data in the scatter plot. The regression line represents the line of best fit. Best fit means that the sum of the squares of the distance between each data point and the line are at a minimum. The formula for the regression line is given as,

$$y' = a + bx$$

where

$$a = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2} \quad b = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

a is the y intercept, and

b is the slope of the line.

The sign of the correlation coefficient and the slope of the regression line will always be the same. The coefficient of determination is the ratio of the explained variation in the data to the total variation and it is usually denoted by  $R^2$ . This coefficient is a measure of the amount of variation in the variable that is described by the regression line. The rest of the

variation or  $1 - R^2$  is undetermined. One method for determining the coefficient of determination is to square the correlation coefficient. The non-parametric test for the correlation coefficient is the Spearman Rank Correlation.

### ***Spearman Rank Correlation***

In the Spearman Rank Correlation, the values of the variables for x and y are expressed in rank order form. This test determines if there is a correlation between the rank order of x and y. The computations for the Pearson coefficient involve ranking each data set and determining the differences between the ranks. These differences are then used to compute  $r_s$ . If both sets of data have the same rank, then  $r_s$  will be +1 or if they are exactly opposite it will be -1. If there is no relationship between the data then the  $r_s$  will be near 0. The formula for  $r_s$  is given by,

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

where d = difference in the ranks, and

n = number of data pairs.

If the test value for  $r_s$  is greater than a given critical value for a 0.05 significance level the null hypothesis is rejected. The value of  $r_s$  translates into a positive correlation between two sets of data when the value is positive. The value of  $r_s$  represents negative correlation when the value is negative.

## **Quantitative Results**

### **Total Mouse Clicks**

The hypothesis was that the enhanced HRI would reduce the number of camera and map clicks. Therefore, the addition of the SES would increase the participants' efficiency.

### ***Camera Clicks***

The purpose of this analysis was to determine if there was any difference in the number of camera clicks dependent upon the order of task presentation. The second purpose was to determine if there was any difference in the number of camera clicks for the original versus the enhanced interface.

In both presentation methods for Tasks A and B, as the participants used the interface they became more comfortable using the camera. With respect to Task A, participants who performed Task A before Task B had more zoom-in and zoom-out clicks for Task A. Also for Task A, participants who performed Task B before Task A had more pan, reset and total clicks. The results also indicate the number of tilt clicks was almost equivalent independent of task order. The results for Task A can be found on the left side of Table 7.1. With respect to Task B, when Task A was performed before Task B there were more tilt, zoom-in, and zoom-out clicks. For Task B before Task A, there were more pan, reset, and total clicks for Task B. The results for Task B can be found on the right side of Table 7.1. These results show that participants who used the SES for their initial task had more total clicks for both tasks. The Task A and B statistical analysis indicated that the relationships between all of these values was insignificant. A table providing the complete results is provided in Appendix I Table 1.

Table 7.1. Tasks A and B Camera Clicks (mean (m), standard deviations (s))

| Click Type | Task A               |                        |                        | Task B                  |                       |                         |
|------------|----------------------|------------------------|------------------------|-------------------------|-----------------------|-------------------------|
|            | Task A before Task B | Task B before Task A   | Total                  | Task A before Task B    | Task B before Task A  | Total                   |
| Pan        | m = 113<br>s = 37.51 | m = 171<br>s = 66.12   | m = 142<br>s = 57.62   | m = 203.6<br>s = 192.52 | m = 308<br>s = 205.44 | m = 294.6<br>s = 188.70 |
| Tilt       | m = 24<br>s = 27.73  | m = 24.33<br>s = 23.18 | m = 24.17<br>s = 22.86 | m = 15.68<br>s = 14.15  | m = 13.5<br>s = 23.69 | m = 15.5<br>s = 20.22   |
| Zoom-In    | m = 5.33<br>s = 0.57 | m = 2.33<br>s = 4.04   | m = 3.83<br>s = 3.06   | m = 3<br>s = 2.65       | m = 1.75<br>s = 3.75  | m = 2<br>s = 3.16       |
| Zoom-Out   | m = 2.68<br>s = 2.52 | m = 1<br>s = 1.73      | m = 1.83<br>s = 21.14  | m = 1.33<br>s = 2.31    | m = 0.75<br>s = 1.5   | m = 1.17<br>s = 1.83    |
| Reset      | m = 2<br>s = 1       | m = 3.68<br>s = 3.79   | m = 2.83<br>s = 2.64   | m = 2<br>s = 1          | m = 3<br>s = 3.37     | m = 2.83<br>s = 2.64    |
| Total      | m = 147<br>s = 64.16 | m = 202.3<br>s = 98.42 | m = 174.6<br>s = 80.25 | m = 225.6<br>s = 203.16 | m = 327<br>s = 199.8  | m = 316.1<br>s = 187.24 |

The comparison of the overall number of camera clicks for Task A versus Task B, indicates that the participants used the camera more for Task B. This was due to a very large number of pan clicks during Task B. The wider hallway in Task B could have attributed to the larger number of pan clicks for this task. The narrow hallway for Task A and the cones being further away from the robot may have attributed to the larger number of tilt, zoom-in, and zoom-out clicks. As indicated by Table 7.2, none of these comparisons between tasks by camera click type was significant.

Table 7.2. Tasks A versus Task B Camera Clicks

| Click Type | Statistic             |
|------------|-----------------------|
| Pan        | $z = -0.73, p = 0.46$ |
| Tilt       | $z = -1.60, p = 0.11$ |
| Zoom-In    | $z = -0.16, p = 0.10$ |
| Zoom-Out   | $z = -1.34, p = 0.17$ |
| Reset      | $z = -1.00, p = 0.31$ |
| Total      | $z = -0.73, p = 0.46$ |

As shown by Table 7.3, participants who performed Task E before Task F had more tilt, zoom-in and reset clicks for Task E. Participants who performed Task F before Task E had more pan, zoom-out and total clicks for Task E. With respect to Task F, participants who performed Task E before Task F had slightly more reset clicks. Participants who performed Task F before Task E had more pan, tilt, zoom-in, zoom-out, and total clicks. Once again, it was shown that participants who used the SES first relied more on the camera for both tasks. The results for Task F are shown on the right side of Table 7.3. All of these results were found to be insignificant as shown in Appendix I Table 2. One possible reason for this result is that participants who performed Task E first may have developed an increased comfort level with the interface, therefore finding it unnecessary to use the camera as much. The large mean value of 476.2 for Task F pan clicks could be attributed to the fact that one participant was an outlier with 1000 pan clicks. When the outlier is removed the mean pan clicks for participants who performed Task F before Task E drops to 343.5 with a standard deviation of 355.76. The overall mean for Task F becomes 256.22 with a standard deviation of 244.76.

Table 7.3. Tasks E and F Camera Clicks (mean (m), standard deviations (s))

|            | Task E                  |                         |                         | Task F                  |                         |                         |
|------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| Click Type | Task E before Task F    | Task F before Task E    | Total                   | Task E before Task F    | Task F before Task E    | Total                   |
| Pan        | m = 311.4<br>s = 315.13 | m = 364.4<br>s = 283.29 | m = 337.9<br>s = 283.87 | m = 186.4<br>s = 105.74 | m = 476.2<br>s = 427.75 | m = 331.3<br>s = 331.09 |
| Tilt       | m = 36<br>s = 37.67     | m = 15.2<br>s = 13.7    | m = 25.60<br>s = 28.88  | m = 34.2<br>s = 50.9    | m = 40.4<br>s = 34.85   | m = 37.30<br>s = 41.26  |
| Zoom-In    | m = 2.4<br>s = 2.3      | m = 2<br>s = 2.92       | m = 2.20<br>s = 2.49    | m = 2<br>s = 2.82       | m = 2.4<br>s = 2.19     | m = 2.20<br>s = 2.39    |
| Zoom-Out   | m = 2<br>s = 3.46       | m = 2.8<br>s = 4.76     | m = 2.40<br>s = 3.95    | m = 0.4<br>s = 0.89     | m = 1.4<br>s = 3.13     | m = 0.90<br>s = 2.23    |
| Reset      | m = 6.4<br>s = 7.73     | m = 4.4<br>s = 1.52     | m = 5.4<br>s = 5.36     | m = 5.8<br>s = 4.2      | m = 5.4<br>s = 2.4      | m = 5.60<br>s = 3.24    |
| Total      | m = 358.2<br>s = 361.9  | m = 388.8<br>s = 292.52 | m = 373.5<br>s = 310.67 | m = 228.8<br>s = 149.13 | m = 525.8<br>s = 446.12 | m = 377.3<br>s = 350.49 |

In a comparison of the number of camera clicks for Task E versus Task F, Table 7.3 demonstrates that Task F had slightly more total camera clicks than Task E. Additionally, Task E had slightly more pan and zoom-out clicks. The Task F camera clicks included one outlier with 1046 total clicks. When the outlier is removed, Task F had an average of 303 total clicks and Task E had an average of 373 total clicks. Therefore, the addition of the SES may have caused the participant to use the camera less to complete the task. As indicated by Table 7.4, the analysis of the individual click types across tasks E and F found no significant relationships.

Table 7.4. Tasks E versus Task F Camera Clicks

| Click Type | Statistic             |
|------------|-----------------------|
| Pan        | $z = -0.08, p = 0.93$ |
| Tilt       | $z = -1.15, p = 0.24$ |
| Zoom-In    | $z = -0.27, p = 0.78$ |
| Zoom-Out   | $z = -1.60, p = 0.10$ |
| Reset      | $z = -0.94, p = 0.34$ |
| Total      | $z = -0.34, p = 0.73$ |

The purpose of the comparison of the number of camera clicks for Task A versus Task E was to determine if the participant used the camera significantly more for the teleoperation tasks since it involved the robot's movement. Table 7.2 and Table 7.3 indicate that Task A had higher zoom-in clicks but Task E had higher pan, tilt, zoom-out, reset, and total clicks. Table 7.5 indicates that none of these results were significant.

Table 7.5. Tasks A versus Task E Camera Clicks

| Click Type | Statistic               |
|------------|-------------------------|
| Pan        | $z = -1.57, p = 0.116$  |
| Tilt       | $z = -0.631, p = 0.528$ |
| Zoom-In    | $z = -0.966, p = 0.334$ |
| Zoom-Out   | $z = -0.378, p = 0.705$ |
| Reset      | $z = -1.084, p = 0.279$ |
| Total      | $z = -1.572, p = 0.116$ |

The purpose of the comparison of the number of camera clicks for Task B versus Task F was to determine if the participant used the camera significantly more for the

teleoperation tasks since it involved the robot's movement. Task B had higher zoom-out clicks but Task F had higher pan, tilt, zoom-in, reset, and total clicks. Table 7.8 indicates that none of these results were significant. Therefore, although it was shown that the more complex task had more camera clicks, there was not a significant difference.

Table 7.6. Tasks B versus Task F Camera Clicks

| Click Type | Statistic               |
|------------|-------------------------|
| Pan        | $z = -1.183, p = 0.237$ |
| Tilt       | $z = -1.521, p = 0.128$ |
| Zoom-In    | $z = -0.677, p = 0.498$ |
| Zoom-Out   | $z = -1.342, p = 0.18$  |
| Reset      | $z = -1.897, p = 0.058$ |
| Total      | $z = -1.352, p = 0.176$ |

### ***Map Clicks***

The purpose of this analysis was to determine if there was any difference in the number of map clicks dependent upon the order of task presentation. The second purpose was to determine if there was any difference in the number of map clicks for the original versus the enhanced interface. The robot's route was essentially identical for Tasks E and F therefore any difference in map clicks are attributed to the interface rather than the task setup. The map was only used for the teleoperation tasks, thus there was only a comparison between Tasks E and F.

Participants who performed Task E before Task F had more move to, map, and total map clicks for Task E as seen in Table 7.7. Participants who performed Task F before Task E had slightly more add icon clicks for the same task. With respect to Task F, participants

who performed Task E before Task F had more move to and total clicks. The number of map and add icon clicks are almost identical, independent of task order presentation. The comparison of the average number of clicks dependent upon task order is located in Table 7.7. The total number of clicks was almost identical between the two tasks. As indicated by Appendix I Table 3, none of these results were found to be significant. The results indicate that participants always used the map more when the original interface was used first. These results suggest that although the SES did not significantly effect the robot teleoperation, there was a learning effect for the second run of Task E. Conversely, there was no learning effect for Task F since the number of map clicks increased for the second run.

Table 7.7. Tasks E and F Map Clicks (mean (m), standard deviation (s))

|            | Task E               |                      |                        | Task F               |                      |                        |
|------------|----------------------|----------------------|------------------------|----------------------|----------------------|------------------------|
| Click Type | Task E before Task F | Task F before Task E | Total                  | Task E before Task F | Task F before Task E | Total                  |
| Move To    | m = 18.0<br>s = 14.3 | m = 13.0<br>s = 7.8  | m = 15.22<br>s = 10.69 | m = 15.8<br>s = 9.1  | m = 13.6<br>s = 2.7  | m = 14.56<br>s = 6.00  |
| Map        | m = 23.8<br>s = 18.9 | m = 20.2<br>s = 9.4  | m = 21.78<br>s = 13.49 | m = 19.5<br>s = 10.1 | m = 19.2<br>s = 7.7  | m = 19.33<br>s = 8.25  |
| Add Icon   | m = 3.0<br>s = 1.8   | m = 3.2<br>s = 1.3   | m = 3.11<br>s = 1.45   | m = 3.5<br>s = 1.9   | m = 3.8<br>s = 2.4   | m = 3.67<br>s = 2.06   |
| Total      | m = 47.8<br>s = 31.6 | m = 39.4<br>s = 17.4 | m = 40.11<br>s = 23.54 | m = 42.5<br>s = 21.3 | m = 38.6<br>s = 12.0 | m = 40.33<br>s = 15.67 |

A comparison of Task E and Task F, shown in Table 7.7, demonstrates that participants required more steps to navigate the robot down the hallway when the SES was

unavailable to aid with landmark determination. As Table 7.8 indicates, none of the comparisons of clicks types across tasks was significant.

Table 7.8. Tasks E versus Task F Map Clicks

| Click Type | Statistic             |
|------------|-----------------------|
| Move To    | $z = 0.0, p = 1.0$    |
| Map        | $z = -0.14, p = 0.88$ |
| Add Icon   | $z = -0.13, p = 0.25$ |
| Total      | $z = -0.21, p = 0.83$ |

A negative correlation exists between the number of camera clicks and the number of map clicks. This result implies that the more the participant used the camera, he/she required fewer steps to teleoperate the robot to the end of the hallway. There was a negative correlation between the number of move to ( $r = -0.822, p = 0.007$ ) and the map clicks ( $r = -0.905, p = 0.001$ ) with the camera zoom-in clicks for Task E. For Task F, there was a negative correlation between the number of move to ( $z = -0.751, p = 0.02$ ) and total map clicks ( $z = -0.716, p = 0.03$ ) and the camera zoom-in clicks. Also for Task F, there was a negative correlation between the total map clicks and the number of camera zoom-in clicks ( $z = -0.786, p = 0.012$ ). This may be because the participant used the camera's zoom to provide a closer view of landmarks and viapoints.

### ***SES Clicks***

The purpose of this analysis is to determine if there is any difference in SES usage dependent upon the order of task presentation. The secondary purpose was to determine if there was any difference in the number of SES clicks for the original versus the enhanced

interface. The hypothesis was that participants may use the SES more and the camera less to accomplish a given task. Additionally, more SES clicks may reduce the number of map clicks. Thus, teleoperation would become more efficient with the addition of the SES.

The SES enhanced the original interface, therefore it was only present for Tasks B and F. Participants who performed Task A before Task B had more scan, reset, zoom-in, zoom-out, tilt up, and tilt down clicks for Task B. Participants who performed Task B before Task A had more pan left, pan right, and total clicks. These results are shown on the left side of Table 7.9.

Participants who performed Task E before Task F had more pan left, tilt down, and total clicks for Task F. Participants who performed Task F before Task E had more scan, reset, zoom-in, zoom-out, pan right, and tilt up clicks. These results are shown on the right side of Table 7.9.

The results for Task B indicate that participants used the SES more (total clicks) to complete the task when the enhanced interface was first. Conversely, for Task F the participants used the SES less to accomplish the task when the enhanced interface was used first. Task B suggests that participants used the SES more when they were not familiar with completing the task without the SES. The other participants had experience completing the task when the SES was absent and may have felt it was not necessary. Appendix I Table 4 indicates that two significant relationships exist. The first result was that participants who performed Task F before Task E had significantly more zoom-out clicks ( $z = -2.2$ ,  $p = 0.025$ ). Additionally, participants who performed Task F before Task E had more pan left clicks ( $z = -2.117$ ,  $p = 0.034$ ). All other results were found insignificant.

Table 7.9. Task B and Task F SES Clicks (mean (m), standard deviations (s))

|            | Task B                 |                         |                         | Task F                  |                         |                         |
|------------|------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| Click Type | Task A before Task B   | Task B before Task A    | Total                   | Task E before Task F    | Task F before Task E    | Total                   |
| Scan       | m = 2.33<br>s = 1.15   | m = 1.75<br>s = 2.22    | m = 2.0<br>s = 1.73     | m = 2.75<br>s = 1.26    | m = 5.2<br>s = 2.77     | m = 4.11<br>s = 2.47    |
| Reset      | m = 2<br>s = 1.0       | m = 1.50<br>s = 2.38    | m = 1.71<br>s = 1.80    | m = 2.75<br>s = 1.26    | m = 5.6<br>s = 2.70     | m = 4.33<br>s = 2.55    |
| Zoom-in    | m = 4.33<br>s = 1.52   | m = 3.25<br>s = 1.26    | m = 3.78<br>s = 1.38    | m = 2.25<br>s = 1.06    | m = 2.8<br>s = 1.64     | m = 2.56<br>s = 1.74    |
| Zoom-out   | m = 2.67<br>s = 1.52   | m = 1.75<br>s = 1.26    | m = 2.14<br>s = 1.35    | m = 1.75<br>s = 2.87    | m = 1.8<br>s = 1.30     | m = 1.79<br>s = 1.99    |
| Pan Left   | m = 7<br>s = 5.29      | m = 39.5<br>s = 76.34   | m = 25.57<br>s = 56.79  | m = 113.2<br>s = 225.17 | m = 43.4<br>s = 55.34   | m = 74.44<br>s = 147.98 |
| Pan Right  | m = 18.3<br>s = 19.6   | m = 112.7<br>s = 135.57 | m = 72.29<br>s = 108.92 | m = 105.2<br>s = 193.4  | m = 131<br>s = 166.57   | m = 119.5<br>s = 167.58 |
| Tilt Up    | m = 4.33<br>s = 2.88   | m = 0.0<br>s = 0.0      | m = 1.86<br>s = 2.85    | m = 17.5<br>s = 37.39   | m = 25.4<br>s = 50.87   | m = 21.89<br>s = 41.29  |
| Tilt Down  | m = 4.33<br>s = 3.79   | m = 0.0<br>s = 0.0      | m = 1.86<br>s = 3.18    | m = 48.75<br>s = 95.51  | m = 4.4<br>s = 9.29     | m = 24.11<br>s = 63.33  |
| Total      | m = 45.33<br>s = 32.02 | m = 160.5<br>s = 172.79 | m = 111.1<br>s = 138.06 | m = 294.2<br>s = 552.59 | m = 219.6<br>s = 227.16 | m = 252.2<br>s = 376.64 |

A comparison of the number of SES clicks for Task B versus Task F indicates that Task F had more clicks for all click types with the exception of zoom-in and zoom-out clicks. This result indicates that during teleoperation, participants used the SES more for the task completion. As indicated by Table 7.10, a comparison of clicks across the tasks found no significant results.

Table 7.10. Tasks B versus Task F SES Clicks

| Click Type | Statistic               |
|------------|-------------------------|
| Scan       | $z = -1.473, p = 0.141$ |
| Reset      | $z = -1.826, p = 0.068$ |
| Zoom-in    | $z = -1.511, p = 0.131$ |
| Zoom-out   | $z = -.730, p = 0.465$  |
| Pan Left   | $z = -.105, p = 0.916$  |
| Pan Right  | $z = -.943, p = 0.345$  |
| Tilt Up    | $z = -.730, p = 0.465$  |
| Tilt Down  | $z = -.730, p = 0.465$  |
| Total      | $z = -.405, p = 0.686$  |

Appendix I Table 5 provides the correlation analysis of the number of SES clicks for Task B versus Task F. This analysis found that a positive correlation exists between the total number of SES clicks for both tasks ( $r = 0.76, p = 0.029$ ). There was also a positive correlation between Task B pan left clicks and Task F zoom-in ( $r = 0.972, p = 0.0$ ), zoom-out ( $r = 0.774, p = 0.024$ ), tilt up ( $r = 0.968, p = 0.0$ ), and total ( $r = 0.749, p = 0.033$ ) clicks. Additionally, there was a positive correlation between total Task B clicks and zoom-in ( $r = 0.737, p = 0.037$ ), zoom-out ( $r = 0.755, p = 0.03$ ), and pan right ( $r = 0.840, p = 0.009$ ) clicks. Therefore, participants who used the SES more for Task B also used it the most for Task F. Appendix I Table 5 only provides significant correlations. Therefore, correlations for the Task F pan left and reset click types were insignificant.

A statistical correlation was also performed between the SES clicks and the camera clicks. It was found that for Task B, there was a negative correlation between the SES zoom-out clicks and the camera zoom-out clicks ( $r = -0.771, p = 0.042$ ). This result suggests that the participant used the SES to replace some of the camera functionality. Additionally, there was a positive correlation between the camera reset clicks and the SES tilt down clicks

for Task F ( $z = -0.73$ ,  $p = 0.017$ ). This result implies that for teleoperation tasks when there was increased use of the camera there was also increased SES use. This could be because it was necessary to use both display devices when the robot was in motion and identifying landmarks. Appendix I Table 6 provides all significant correlations between the camera click types and the SES click types. The remainder of the SES and camera click types were insignificant.

A statistical correlation was performed between the SES clicks and the map clicks. It was found that no significant correlations existed between the SES clicks and the map clicks.

### **Task Scores**

The hypothesis was that the enhanced Human-Robot interface along with the learning effect would increase task scores. The purpose of this analysis is to determine if there is any difference in task scores dependent upon the order of task presentation. The secondary purpose was to determine if there was any difference in task scores for the original versus the enhanced interface.

In Tasks A and B, the robot was stationary and the participants were required to find the robot and mark the location on a landmark map. Additionally, the participants were required to mark all locations of landmarks around the robot on the map. Participants who performed Task A before Task B scored higher for cone placement and cone color for Task A. Participants who performed Task B before Task A scored higher for the driving directions and slightly higher for the overall score on Task A, as shown in Table 7.11. The robot placement and robot orientation scores were the same independent of task presentation.

Participants who performed Task A before Task B scored higher on robot orientation, cone color, and driving directions during Task B. Participants who performed Task B before Task A scored higher for robot placement, cone placement, and the overall score, as shown in Table 7.11. The relationship between tasks scores and task presentation were all found to be insignificant and are presented in Appendix I Table 7. These results contradict the learning theory because participants scored higher on Task B even though it was performed first. Therefore, the addition of the SES may have indeed increased the task score. The comparison of these results can be found in Table 7.11. As previously stated, none of these results were found to be statistically significant. One possible source for this disparity, other than the addition of the SES, is that at the task B location, there was an electric field that created more error in the compass. This error may have caused participants to make heading errors.

Table 7.11. Tasks A and B Scores (mean (m), standard deviations (s))

| Sub-Score    | Task A               |                        |                        | Task B                 |                        |                        |
|--------------|----------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
|              | Task A before Task B | Task B before Task A   | Total                  | Task A before Task B   | Task B before Task A   | Total                  |
| Robot Plcmt  | m = 100<br>s = 0     | m = 100<br>s = 0       | m = 100<br>s = 0.0     | m = 66.67<br>s = 57.74 | m = 100<br>s = 0       | m = 87.5<br>s = 35.36  |
| Robot Orient | m = 100<br>s = 0     | m = 100<br>s = 0       | m = 100<br>s = 0.0     | m = 100<br>s = 0.0     | m = 60<br>s = 54.77    | m = 75.00<br>s = 46.29 |
| Cone Plcmt   | m = 90<br>s = 14.9   | m = 88.89<br>s = 19.25 | m = 89.58<br>s = 15.27 | m = 41.67<br>s = 52.04 | m = 83.33<br>s = 15.59 | m = 67.71<br>s = 37.12 |
| Cone Color   | m = 100<br>s = 0     | m = 88.89<br>s = 19.25 | m = 95.83<br>s = 11.79 | m = 91.67<br>s = 14.43 | m = 83.33<br>s = 15.59 | m = 86.46<br>s = 14.73 |

Table 7.11, continued

|               |           |           |           |           |           |           |
|---------------|-----------|-----------|-----------|-----------|-----------|-----------|
| Driving       | m = 80    | m = 93.33 | m = 85    | m = 100   | m = 96    | m = 97.14 |
| Direct        | s = 44.72 | s = 11.55 | s = 35.05 | s = 0.0   | s = 8.94  | s = 7.56  |
| Overall Score | m = 94.07 | m = 93.83 | m = 93.98 | m = 75.77 | m = 84.78 | m = 81.4  |
|               | s = 7.68  | s = 7.71  | s = 7.12  | s = 21.72 | s = 8.94  | s = 14.16 |

In a comparison of Task A versus Task B, participants had higher scores for robot placement, robot orientation, cone placement, cone color, and overall score for Task A. Therefore, the SES actually decreased the task scores on all but the driving directions. Table 7.12 presents the comparison of task scores between tasks A and B, none of these results were significant.

Table 7.12. Tasks A versus Task B Scores

| Sub-Score         | Statistic             |
|-------------------|-----------------------|
| Robot Placement   | $z = -1.0, p = 0.31$  |
| Robot Orientation | $z = -1.4, p = 0.15$  |
| Cone Placement    | $z = -0.94, p = 0.34$ |
| Cone Color        | $z = -1.08, p = 0.27$ |
| Driving Direction | $z = -1.089, p = 1.0$ |
| Overall Score     | $z = -1.78, p = 0.07$ |

In the teleoperation tasks E and F, there were different sub-tasks than those for in Tasks A and B. The sub-tasks only included cone placement and cone color, as seen in Table 7.13. Participants who performed Task F before Task E had higher cone placement, cone color, and an overall score for both tasks. These results show that Task E demonstrates learning similarly to Task A. The result for Task F indicates that the SES may

have attributed to the increased scores. Appendix I Table 8 indicates that none of these relationships were significant.

Table 7.13. Tasks E and F Scores (mean (m), standard deviations (s))

| Sub-Score     | Task E                 |                        |                        | Task F                 |                        |                        |
|---------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
|               | Task E before Task F   | Task F before Task E   | Total                  | Task E before Task F   | Task F before Task E   | Total                  |
| Cone Plcmt    | m = 65.71<br>s = 19.17 | m = 68.57<br>s = 39.64 | m = 67.14<br>s = 29.39 | m = 72.86<br>s = 20.45 | m = 84.29<br>s = 16.29 | m = 78.57<br>s = 18.44 |
| Cone Color    | m = 91.43<br>s = 12.78 | m = 94.29<br>s = 12.78 | m = 92.85<br>s = 12.14 | m = 82.86<br>s = 23.47 | m = 85.71<br>s = 17.50 | m = 84.29<br>s = 19.58 |
| Overall Score | m = 78.57<br>s = 13.36 | m = 81.43<br>s = 26.05 | m = 80<br>s = 18.57    | m = 77.86<br>s = 14.81 | m = 85.00<br>s = 11.95 | m = 81.43<br>s = 13.24 |

In a comparison of Task E versus Task F, Task E had a higher cone color score while Task F had a higher cone placement score and slightly higher overall score. The SES may have improved the resulting task score, but not significantly. Table 7.14 presents the comparison between tasks. The comparison of cone color score found a slightly insignificant result while the remaining relationships were clearly insignificant.

Table 7.14. Tasks E versus Task F Scores

| Sub-Score      | Statistic             |
|----------------|-----------------------|
| Cone Placement | $z = -1.26, p = 0.20$ |
| Cone Color     | $z = -1.76, p = 0.07$ |
| Overall Score  | $z = -0.17, p = 0.85$ |

The correlation analysis between the task score and the number of camera clicks indicates that an over dependence on the camera actually had a negative effect on the task score. Appendix I Table 9 details the results of this analysis. For Task A, there was a negative correlation between the driving directions score and the pan ( $r = -0.859$ ,  $p = 0.029$ ), reset ( $r = -0.959$ ,  $p = 0.002$ ), and total number of camera clicks ( $r = -0.826$ ,  $p = 0.043$ ). For Task B, there was a negative correlation between the driving directions score and the number of tilt ( $r = -0.987$ ,  $p = 0.0$ ) and reset ( $r = -0.956$ ,  $p = 0.003$ ) clicks. Also for Task B, there was a negative correlation between the robot placement score and the number of zoom-out clicks ( $r = -0.764$ ,  $p = 0.046$ ). For Task E, there was a negative correlation between the overall score and the number of zoom-out clicks ( $r = -0.748$ ,  $p = 0.013$ ). Finally, for Task F there was positive correlation between the number of reset clicks and the cone placement score ( $r = 0.717$ ,  $p = 0.02$ ). No significant correlations existed between the task score and the number of map clicks.

The correlation performed between the task score and the number of SES clicks showed that the more the participant used the SES, the lower the score. This result is in direct contradiction with the hypothesis that will be addressed in the next chapter. The SES was only present on the enhanced interface, therefore Appendix I Table 10 only contains correlation analysis for Tasks B and F. For Task B, there was a negative correlation between the cone color score and the pan left ( $r = -0.679$ ,  $p = 0.064$ ) and total clicks ( $r = -0.844$ ,  $p = 0.008$ ). Also for Task B, there was a negative correlation between the robot orientation score and the pan right ( $r = -0.917$ ,  $p = 0.001$ ) and total ( $r = -0.810$ ,  $p = 0.015$ ) clicks. For Task F, there was a negative correlation between the SES tilt up clicks and the cone color score ( $r = -0.675$ ,  $p = 0.032$ ). These results indicate that the task score decreased as use of the SES increased.

### Task Completion Times

The hypothesis was that the enhanced HRI along with the learning effect would decrease the task completion times. The purpose of this analysis is to determine if there is any difference in task completion time dependent upon the order of task presentation. The secondary purpose was to determine if there was any difference in task completion time for the original versus the enhanced interface.

The results demonstrated that participants who performed Task B before Task A required less time to find the robot, drive the robot, and complete the tasks for Task A. Participants who performed Task A before Task B required less time to find the robot, drive the robot, as well as complete Task B. These results indicate a definite learning effect since the second run of the task was much faster. This reduction was most likely due to the participant becoming familiar and more confident with the interface. As indicated by Appendix I Table 11, none of these results were significant.

Table 7.15. Tasks A and B Completion Times (mean (m), standard deviations (s))

|                 | Task A                 |                        |                        | Task B                 |                        |                        |
|-----------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| Sub-Task        | Task A before Task B   | Task B before Task A   | Total                  | Task A before Task B   | Task B before Task A   | Total                  |
| Find the Robot  | m = 08:51<br>s = 03:48 | m = 05:19<br>s = 00:53 | m = 07:32<br>s = 03:26 | m = 07:48<br>s = 00:56 | m = 14:12<br>s = 13:55 | m = 11:48<br>s = 11:43 |
| Drive the Robot | m = 02:20<br>s = 01:23 | m = 00:52<br>s = 00:22 | m = 01:47<br>s = 01:18 | m = 00:50<br>s = 00:48 | m = 02:00<br>s = 00:43 | m = 01:34<br>s = 00:59 |
| Overall Time    | m = 11:11<br>s = 03:52 | m = 06:12<br>s = 01:14 | m = 09:19<br>s = 03:57 | m = 08:38<br>s = 01:09 | m = 16:13<br>s = 14:03 | m = 13:22<br>s = 12:03 |

In a comparison of Task A versus Task B, Task B had a lower drive the robot time but Task A had lower find the robot and overall times. This result contradicts the hypothesis that the enhanced interface would reduce the task completion time. Table 7.16 indicates that none of these comparisons were statistically significant.

Table 7.16. Tasks A versus Task B Completion Times

| Sub-Task        | Statistic             |
|-----------------|-----------------------|
| Find the Robot  | $z = -0.94, p = 0.34$ |
| Drive the Robot | $z = -0.10, p = 0.91$ |
| Overall Time    | $z = -0.73, p = 0.46$ |

Since the teleoperation tasks did not include the same sub-tasks as Tasks A and B, there was only an overall time for Tasks E and F. The results for Task E and Task F were very similar to those found for Tasks A and B. The participants who performed Task F before Task E greatly reduced their overall Task E time, on average by nine minutes. Participants who performed Task E before Task F reduced their average Task F time by six minutes. Table 7.17 provides the comparison of the task completion times dependent upon task order.

Table 7.17. Tasks E and Task F Completion Times (mean (m), standard deviations (s))

|                 | Task E                     |                            |                        | Task F                     |                            |                        |
|-----------------|----------------------------|----------------------------|------------------------|----------------------------|----------------------------|------------------------|
|                 | Task E<br>before<br>Task F | Task F<br>before<br>Task E | Total                  | Task E<br>before<br>Task F | Task F<br>before<br>Task E | Total                  |
| Overall<br>Time | m = 22:36<br>s = 01:27     | m = 13:16<br>s = 03:35     | m = 17:56<br>s = 05:33 | m = 17:14<br>s = 04:35     | m = 23:18<br>s = 06:54     | m = 20:16<br>s = 06:23 |

The analysis across tasks showed that the completion time dependent on task presentation was significant for Task E. Participants who performed Task E before Task F averaged completion times that were 9 minutes longer for Task E ( $r = -2.61$ ,  $p = 0.0009$ ). The statistical analysis results for Tasks E and F are given in Table 7.18.

Table 7.18. Tasks E and F Completion Times (Statistics)

|                 | Task E                      | Task F                      |
|-----------------|-----------------------------|-----------------------------|
| Overall<br>Time | $z = -2.61$<br>$p = 0.0009$ | $z = -1.567$<br>$p = 0.117$ |

In a comparison of Task E versus Task F, the task completion time was longer for Task F by an average of 2 minutes. This result indicates that the addition of the SES did not improve the task time. This difference in task time was not significant ( $z = -0.56$ ,  $p = 0.57$ ).

The correlation between the number of camera clicks and the task completion time demonstrated that the more the camera was used, more was time required to complete the task. As presented in Appendix I Table 12, there was a positive correlation between the number of zoom-out clicks and the completion time ( $r = 0.832$ ,  $p = 0.04$ ) for Task A. For Task B, there was a positive correlation between the number of tilt ( $r = 0.801$ ,  $p = 0.03$ ) and

reset ( $r = 0.94$ ,  $p = 0.002$ ) clicks with the find the robot time. Also for Task B, there was a positive correlation between the driving directions time and the number of pan ( $r = 0.84$ ,  $p = 0.016$ ) and total clicks ( $r = 0.861$ ,  $p = 0.013$ ). Finally for Task B, there was a positive correlation between total task time and the number of tilt ( $r = 0.79$ ,  $p = 0.033$ ) and reset ( $r = 0.94$ ,  $p = 0.002$ ) clicks. There were no significant correlations for Task E. For Task F, there was a positive correlation between the total number of clicks and the completion time ( $r = 0.713$ ,  $p = 0.021$ ). There are no significant correlations between the number of zoom-in clicks and the task times. No significant correlations existed between task completion times and the number of map clicks.

A correlation between completion times and SES clicks found a negative correlation between the find the robot time ( $r = -0.72$ ,  $p = 0.043$ ) and the total task time ( $r = -0.717$ ,  $p = 0.045$ ) with the number of SES zoom-out clicks for Task B. None of the other SES click type correlations were significant. This result implies that the use of the SES actually did reduce the task completion time for Task B. There were no significant results for Task F. These results can be found in Appendix I Table 13.

The correlation between the task completion time and task score only found significant results for Task B, which are provided in Appendix I Table 14. There was a negative correlation between the driving directions score and the find the robot ( $r = -0.99$ ,  $p = 0.0$ ) and total task time ( $r = -0.99$ ,  $p = 0.0$ ) for Task B. There were no significant correlations for the remainder of the task scores (robot placement, robot orientation, cone placement, cone color, and overall score). The negative correlation between the score and task time suggests a reduction in task score as task completion time became longer.

### **Multiple Resources Questionnaire (MRQ)**

The hypothesis is that the enhanced HRI and learning should result in the participant using fewer resources to complete tasks. The purpose of this analysis is to determine if there is any difference in responses to the MRQ dependent upon the order of task presentation. The secondary purpose was to determine if there was any difference in responses to the MRQ for the original versus the enhanced interface. The actual MRQ questionnaire is located in Appendix B. The rating scale for the MRQ was 0 to 5.

Participants who performed Task A before Task B had higher responses for short-term memory, spatial attentive, spatial categorical, spatial positional, visual lexical, visual temporal, and overall resources for Task A, as provided in Table 7.19. Participants who performed Task B before Task A provided higher responses for manual, spatial emergent, and spatial quantitative processes for Task A. These results imply that Task A potentially required fewer resources the more the task was performed. Participants who performed Task A before Task B provided higher responses for spatial categorical, spatial quantitative, and visual lexical processes for Task B. Participants who performed Task B before Task A indicated higher responses for spatial attentive, spatial emergent, and visual temporal resources. All other Task B resources were rated identically independent of task presentation. Additionally, the total resources mean was equivalent therefore, Task B had the same demand on resources independent of the task presentation. The results for the Task A and B comparison are provided in Table 7.19. As indicated by Appendix I Table 15, none of these results were significant.

Table 7.19. Tasks A and Task B MRQ (mean (m), standard deviations (s))

|                      | Task A               |                      |                      | Task B               |                      |                      |
|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Sub-Process          | Task A before Task B | Task B before Task A | Total                | Task A before Task B | Task B before Task A | Total                |
| Manual               | m = 2.4<br>s = 0.55  | m = 2.67<br>s = 2.08 | m = 2.5<br>s = 1.20  | m = 2.00<br>s = 0.00 | m = 2.00<br>s = 0.00 | m = 2.38<br>s = 1.06 |
| Short-term memory    | m = 3.4<br>s = 0.89  | m = 2.67<br>s = 0.58 | m = 3.13<br>s = 0.83 | m = 3.00<br>s = 0.00 | m = 3.00<br>s = 0.00 | m = 3.00<br>s = 0.00 |
| Spatial attentive    | m = 4.2<br>s = 0.84  | m = 3.33<br>s = 1.53 | m = 3.88<br>s = 1.13 | m = 3.33<br>s = 0.58 | m = 4.00<br>s = 0.00 | m = 3.75<br>s = 0.71 |
| Spatial categorical  | m = 4.2<br>s = 0.45  | m = 3.00<br>s = 1.73 | m = 3.75<br>s = 1.16 | m = 3.33<br>s = 0.58 | m = 3.00<br>s = 0.00 | m = 3.38<br>s = 0.74 |
| Spatial emergent     | m = 0.60<br>s = 0.89 | m = 2.00<br>s = 0.00 | m = 1.13<br>s = 0.99 | m = 2.67<br>s = 1.15 | m = 3.33<br>s = 1.15 | m = 2.75<br>s = 1.04 |
| Spatial positional   | m = 3.8<br>s = 2.17  | m = 2.33<br>s = 2.52 | m = 3.25<br>s = 2.25 | m = 1.33<br>s = 2.31 | m = 1.33<br>s = 2.31 | m = 1.63<br>s = 2.26 |
| Spatial quantitative | m = 1.8<br>s = 2.05  | m = 2.00<br>s = 0.00 | m = 1.88<br>s = 1.55 | m = 1.67<br>s = 0.58 | m = 1.33<br>s = 1.15 | m = 1.63<br>s = 0.74 |
| Visual lexical       | m = 2.4<br>s = 1.52  | m = 1.33<br>s = 1.15 | m = 2.00<br>s = 1.41 | m = 1.67<br>s = 0.58 | m = 0.67<br>s = 1.15 | m = 1.13<br>s = 0.99 |
| Visual temporal      | m = 2.2<br>s = 1.64  | m = 0.67<br>s = 0.58 | m = 1.63<br>s = 1.51 | m = 0.67<br>s = 0.58 | m = 1.00<br>s = 1.00 | m = 0.88<br>s = 0.64 |
| Overall Resources    | m = 2.78<br>s = 0.89 | m = 2.22<br>s = 0.77 | m = 2.57<br>s = 0.84 | m = 2.19<br>s = 0.28 | m = 2.19<br>s = 0.39 | m = 2.28<br>s = 0.43 |

With respect to Task A versus Task B, there is no significant difference between the resource ratings across tasks as indicated by Table 7.20. The results did indicate that the overall resources for Task A were slightly higher than for Task B, which indicates the SES may have caused a slight difference as shown by Table 7.19. This difference may be due to Task B including the SES while Task A did not. Task A had higher responses for manual,

short-term memory, spatial attentive, spatial categorical, spatial positional, spatial quantitative, visual lexical, visual temporal, and overall resources. Task B was rated higher for only the spatial emergent resources. This difference could be attributed to the fact that those participants who worked with the SES first had to perform more spatial judgments when the SES was not present.

Table 7.20. Tasks A versus Task B MRQ

| Sub-Process          | Statistic             |
|----------------------|-----------------------|
| Manual               | $z = -1.0, p = 0.31$  |
| Short-term memory    | $z = -1.0, p = 0.31$  |
| Spatial attentive    | $z = 0.0, p = 1.0$    |
| Spatial categorical  | $z = -0.37, p = 0.70$ |
| Spatial emergent     | $z = -1.34, p = 0.18$ |
| Spatial positional   | $z = -0.92, p = 0.35$ |
| Spatial quantitative | $z = -0.27, p = 0.78$ |
| Visual lexical       | $z = -0.27, p = 1.0$  |
| Visual temporal      | $z = -1.0, p = 0.31$  |
| Overall Resources    | $z = 0.0, p = 1.0$    |

In the comparison of Tasks E and F, participants who performed Task E before Task F had higher responses for the spatial attentive and visual temporal processes as presented in Table 7.21. Participants who performed Task F before Task E had higher responses for short-term memory, spatial categorical, spatial emergent, spatial positional, spatial quantitative, visual lexical and overall resources. All other responses were equivalent independent of task order. These results are in direct contradiction to the belief that the second run of a task would require lower multiple resources demand. Therefore, this result may suggest that the presence of the SES for the initial run of the task caused the

participants to use more resources when it was removed during the second run of the task. The results for Task E are located on the left side of Table 7.21. Participants who performed Task E before Task F had higher responses for spatial attentive, spatial positional, spatial quantitative, visual lexical, visual temporal and the overall resources. Participants who performed Task F before Task E had a higher demand on the short-term memory and spatial emergent resources. All other responses were equivalent independent of task order. These results are provided in Table 7.21. These results indicate the second run of the task caused a higher demand on the overall resources. Finally, Appendix I Table 16 indicates that none of these results was significant.

Table 7.21. Task E and Task F MRQ (mean (m), standard deviations (s))

|                        | Task E                     |                            |                      | Task F                     |                            |                      |
|------------------------|----------------------------|----------------------------|----------------------|----------------------------|----------------------------|----------------------|
| Sub-Process            | Task E<br>before<br>Task F | Task F<br>before<br>Task E | Total                | Task E<br>before<br>Task F | Task F<br>before<br>Task E | Total                |
| Manual                 | m = 2.00<br>s = 0.71       | m = 2.00<br>s = 0.00       | m = 2.00<br>s = 0.47 | m = 2.00<br>s = 0.00       | m = 2.00<br>s = 0.71       | m = 2.00<br>s = 0.47 |
| Short-term<br>memory   | m = 2.80<br>s = 0.45       | m = 3.00<br>s = 0.71       | m = 2.90<br>s = 0.57 | m = 3.00<br>s = 0.71       | m = 3.20<br>s = 0.45       | m = 3.10<br>s = 0.57 |
| Spatial<br>attentive   | m = 3.60<br>s = 0.89       | m = 3.00<br>s = 0.71       | m = 3.30<br>s = 0.82 | m = 3.60<br>s = 0.55       | m = 3.20<br>s = 0.84       | m = 3.40<br>s = 0.70 |
| Spatial<br>categorical | m = 3.00<br>s = 1.87       | m = 3.60<br>s = 0.89       | m = 3.30<br>s = 1.42 | m = 3.40<br>s = 0.55       | m = 3.40<br>s = 1.34       | m = 3.40<br>s = 0.97 |
| Spatial<br>emergent    | m = 1.40<br>s = 1.67       | m = 2.00<br>s = 2.00       | m = 1.70<br>s = 1.77 | m = 2.00<br>s = 1.41       | m = 2.80<br>s = 1.10       | m = 2.40<br>s = 1.26 |
| Spatial<br>positional  | m = 1.60<br>s = 1.67       | m = 2.40<br>s = 1.67       | m = 2.00<br>s = 1.63 | m = 2.80<br>s = 1.79       | m = 1.60<br>s = 1.67       | m = 2.20<br>s = 1.75 |

Table 7.1, continued

|                      |                      |                      |                      |                      |                      |                      |
|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Spatial quantitative | m = 1.40<br>s = 1.67 | m = 3.20<br>s = 1.10 | m = 2.30<br>s = 1.64 | m = 1.80<br>s = 1.48 | m = 1.60<br>s = 1.67 | m = 1.70<br>s = 1.49 |
| Visual lexical       | m = 1.60<br>s = 1.52 | m = 2.00<br>s = 1.41 | m = 1.80<br>s = 1.40 | m = 2.00<br>s = 1.87 | m = 1.4<br>s = 1.67  | m = 1.70<br>s = 1.70 |
| Visual temporal      | m = 1.80<br>s = 2.05 | m = 1.00<br>s = 1.00 | m = 1.40<br>s = 1.58 | m = 2.20<br>s = 1.10 | m = 2.00<br>s = 1.63 | m = 2.11<br>s = 1.27 |
| Overall Resources    | m = 2.13<br>s = 0.55 | m = 2.47<br>s = 0.56 | m = 2.30<br>s = 0.55 | m = 2.53<br>s = 0.58 | m = 2.35<br>s = 0.67 | m = 2.44<br>s = 0.60 |

In the comparison of Task E versus Task F, Table 7.22 indicates that there are no significant differences in the MRQ processes. Task E had higher spatial quantitative and visual lexical processes. Task F had higher short-term memory, spatial attentive, spatial categorical, spatial emergent, spatial positional, visual temporal, and overall ratings. The manual processes were the same for both tasks. Therefore, the conclusion is that the SES did not assist in reducing the multiple resources.

Table 7.22. Task E versus Task F MRQ

| Sub-Process          | Statistic              |
|----------------------|------------------------|
| Manual               | $z = 0.0, p = 1.0$     |
| Short-term memory    | $z = -1.41, p = 0.15$  |
| Spatial attentive    | $z = -0.57, p = 0.56$  |
| Spatial categorical  | $z = -0.18, p = 0.85$  |
| Spatial emergent     | $z = -0.95, p = 0.33$  |
| Spatial positional   | $z = -0.272, p = 0.78$ |
| Spatial quantitative | $z = -0.75, p = 0.45$  |

Table 7.22, continued

|                   |                       |
|-------------------|-----------------------|
| Visual lexical    | $z = -0.13, p = 0.89$ |
| Visual temporal   | $z = -1.51, p = 0.13$ |
| Overall Resources | $z = -0.89, p = 0.37$ |

When comparing Task A and Task E, no significant difference was found in the number of resources participants' used as shown in Table 7.23. A similar comparison between Task B and Task F found no significant difference in the multiple resources required between the tasks employing the SES. In both cases, the data suggests that the participants appear to have used fewer resources when driving the robot. This is an unexpected and contradictory result.

Table 7.23. Task Comparison MRQ

| Sub-Process          | A versus E            | B versus F             |
|----------------------|-----------------------|------------------------|
| Manual               | $z = -1.41, p = 0.15$ | $z = -1.34, p = 0.17$  |
| Short-term memory    | $z = -0.44, p = 0.15$ | $z = -1.0, p = 0.31$   |
| Spatial attentive    | $z = -1.29, p = 0.19$ | $z = -1.632, p = 0.10$ |
| Spatial categorical  | $z = -0.92, p = 0.35$ | $z = -0.33, p = 0.73$  |
| Spatial emergent     | $z = -1.29, p = 0.19$ | $z = -1.73, p = 0.08$  |
| Spatial positional   | $z = -0.96, p = 0.33$ | $z = -0.17, p = 0.86$  |
| Spatial quantitative | $z = 0.0, p = 1.0$    | $z = -0.57, p = 0.56$  |
| Visual lexical       | $z = -0.37, p = 0.70$ | $z = -0.27, p = 0.78$  |
| Visual temporal      | $z = -0.14, p = 0.88$ | $z = -1.46, p = 0.14$  |
| Overall Resources    | $z = -0.56, p = 0.57$ | $z = -0.67, p = 0.49$  |

Appendix I Table 17 provides the correlation between the MRQ and the number of camera clicks. There were no significant correlations for Task A. For Task B, a negative

correlation existed between the spatial quantitative resources and the zoom-in ( $r = -0.862$ ,  $p = 0.013$ ), zoom-out ( $r = -0.85$ ,  $p = 0.014$ ), and reset clicks ( $r = -0.95$ ,  $p = 0.001$ ). There was also a negative correlation between the number of camera tilt clicks and the visual temporal resources ( $r = -0.835$ ,  $p = 0.019$ ) for Task B. For Task E, there was a negative correlation between total clicks and the visual temporal resource ( $r = -0.631$ ,  $p = 0.05$ ). Also for Task E, there was a negative correlation between the overall resource and the number of zoom-in clicks ( $r = -0.668$ ,  $p = 0.035$ ). Finally, Task E showed a negative correlation between the total camera clicks and the manual process ( $r = -0.647$ ,  $p = 0.043$ ). For Task F, there was a negative correlation between the spatial attentive resources with the total clicks ( $r = -0.667$ ,  $p = 0.035$ ) as well as for the pan clicks ( $r = -0.679$ ,  $p = 0.031$ ). A negative correlation also existed between the visual temporal resources and the number of pan ( $r = -0.818$ ,  $p = 0.007$ ), reset ( $r = -0.693$ ,  $p = 0.039$ ) and total ( $r = -0.799$ ,  $p = 0.01$ ) clicks. There was a negative correlation between the number of zoom-in clicks and the spatial quantitative resource ( $r = -0.664$ ,  $p = 0.036$ ) for Task F. Finally, there was a positive correlation between the number of zoom-out clicks and the manual process ( $r = 0.739$ ,  $p = 0.015$ ). The remaining click types and resources did not exhibit significant correlations. These results suggest that the more the camera was used, the less the demand on the specified multiple resources. The single positive correlation may suggest that zooming the camera excessively increases a demand on the manual processes.

Appendix I Table 18 provides the correlation between the MRQ and the number of map clicks. For Task E, there was a positive correlation between the manual process and the number of move to point ( $r = 0.678$ ,  $p = 0.045$ ), map ( $r = 0.704$ ,  $p = 0.034$ ) and total clicks ( $r = 0.68$ ,  $p = 0.044$ ). There was also positive correlations between the visual temporal resources and the number of move to point ( $r = 0.75$ ,  $p = 0.02$ ), map ( $r = 0.691$ ,  $p = 0.039$ ),

add icon ( $r = 0.692$ ,  $p = 0.039$ ) and total ( $r = 0.694$ ,  $p = 0.038$ ) clicks. These results signify that there is more movement of hands and fingers with higher camera clicks for Task E therefore a higher demand on the manual processes. Also for Task E, there was a higher judgement of time intervals using the sense of vision since the robot was moving and thus a higher visual temporal demand when using the camera more. There were no correlations for Task F between the MRQ scores and the number of map clicks. Additionally, the remaining relationships were insignificant.

Appendix I Table 19 provides the correlation between the MRQ and the number of SES clicks. For Task B, there was a positive correlation between the number of zoom-out clicks and the spatial quantitative resources ( $r = 0.861$ ,  $p = 0.006$ ). There was a positive correlation for Task B between the number of pan left clicks and the spatial positional resources ( $r = 0.772$ ,  $p = 0.025$ ) as well as a positive correlation between the overall resources and the number of pan left clicks ( $r = 0.764$ ,  $p = 0.027$ ). The analysis for Task F found a negative correlation between the number of zoom-in clicks and the visual temporal resources ( $r = -0.653$ ,  $p = 0.041$ ). There were positive correlations between the spatial quantitative resources in Task F with the zoom-out ( $r = 0.69$ ,  $p = 0.027$ ), pan left ( $r = 0.717$ ,  $p = 0.02$ ), pan right ( $r = 0.878$ ,  $p = 0.0001$ ), as well as total clicks ( $r = 0.807$ ,  $p = 0.005$ ). Finally, a negative correlation existed between the number of tilt down clicks and the spatial emergent process ( $r = -0.687$ ,  $p = 0.028$ ). The remaining comparisons did not exhibit significant correlations. These results signify that there may have been a higher demand on multiple resources when there was a greater number of SES clicks. This is a contradiction to the hypothesis that the addition of the SES may reduce the demand on multiple resources.

Appendix I Table 20 provides the correlation between the MRQ and the task scores. Negative correlations existed between the driving directions score versus the overall

resources ( $r = -0.72$ ,  $p = 0.04$ ) as well as the visual lexical process and the overall score ( $r = -0.74$ ,  $p = 0.04$ ) for Task A. For Task B, there was a positive correlation between the driving directions score and the spatial quantitative ( $r = 0.88$ ,  $p = 0.009$ ) as well as the visual temporal resources ( $r = 0.76$ ,  $p = 0.046$ ). There was a negative correlation for Task F between the overall resources and the overall score ( $r = -0.77$ ,  $p = 0.009$ ). There were no significant correlations for Task E score and the MRQ. The remainder of the comparisons did not exhibit any significant correlations. These results indicate that there is a reduction in the task score when there is an increased demand on multiple resources.

Finally, Appendix I Table 21 provides the correlation analysis between the MRQ and the task completion times. These results show that for Task A there is a negative correlation between the total task time and the spatial emergent process ( $r = -0.754$ ,  $p = 0.031$ ). There is a negative correlation between the spatial quantitative process and the overall task time ( $r = -0.873$ ,  $p = 0.005$ ) as well as the find the robot time ( $r = -0.893$ ,  $p = 0.003$ ) for Task B. There were no significant results for Tasks E and F. The results of Task A and B signify that there was a higher demand on multiple resources for a shorter task completion time.

### **NASA-TLX Workload Rating**

The hypothesis is that the addition of the SES to the HRI would reduce the participants' perceived workload. The purpose of this analysis is to determine if perceived workload is dependent upon the task order. The secondary purpose was to determine if perceived workload is affected by the introduction of the SES to the HRI. The actual NASA-TLX questionnaire is located in Appendix A. The rating scale was 0 to 100. The overall workload rating was determined by taking an average of all of the sub-scale responses.

As indicated by Table 7.24, participants who performed Task A before Task B rated the amount of necessary thinking, task difficulty, physical demand, time required, time pressure, performance satisfaction, mental effort, frustration level, stress level and overall ratings for Task A higher. Participants who performed Task B before Task A rated their goal achievement and physical effort higher for Task A. Participants who performed Task A before Task B rated the task difficulty, time required, performance satisfaction, and mental effort higher for Task B. Participants who performed Task B before Task A had higher amounts of necessary thinking, time pressure, goal achievement, physical effort, frustration level, stress level, and overall ratings for Task B. The physical demand was equivalent of Task B independent of task presentation. These results indicate that there was a higher perceived workload the first time a task was performed. This is intuitive as the task should become easier the as the user performs the task more frequently. Appendix I Table 22 indicates that the only significant result related to task order compared to perceived workload was for Task A, in which the mental effort was significantly higher when Task A was performed before Task B ( $z = -2.23$ ,  $p = 0.025$ ).

Table 7.24. Tasks A and Task B NASA-TLX (mean (m), standard deviations (s))

| Sub-Rating         | Task A                |                      |                      | Task B               |                      |                      |
|--------------------|-----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                    | Task A before Task B  | Task B before Task A | Total                | Task A before Task B | Task B before Task A | Total                |
| Necessary thinking | m = 56.2<br>s = 18.47 | m = 43.0<br>s = 49.8 | m = 51.2<br>s = 30.8 | m = 49<br>s = 15.5   | m = 62.8<br>s = 32.9 | m = 57.6<br>s = 27.2 |
| Task difficulty    | m = 35.0<br>s = 25.33 | m = 13.3<br>s = 7.02 | m = 26.8<br>s = 22.5 | m = 33.6<br>s = 19.8 | m = 26.4<br>s = 30.9 | m = 29.1<br>s = 25.9 |

Table 7.24, continued

|                          |                        |                      |                       |                      |                       |                       |
|--------------------------|------------------------|----------------------|-----------------------|----------------------|-----------------------|-----------------------|
| Physical demand          | m = 1.60<br>s = 3.58   | m = 0.0<br>s = 0.0   | m = 1.00<br>s = 2.83  | m = 0.0<br>s = 0.0   | m = 0.0<br>s = 0.0    | m = 0.0<br>s = 0.0    |
| Time required            | m = 26.00<br>s = 17.16 | m = 24.6<br>s = 29.9 | m = 25.5<br>s = 20.6  | m = 23<br>s = 12.17  | m = 21.2<br>s = 25.63 | m = 21.8<br>s = 20.46 |
| Time pressure            | m = 24.6<br>s = 39.07  | m = 0.33<br>s = 0.58 | m = 15.5<br>s = 32.09 | m = 2.67<br>s = 3.06 | m = 7.4<br>s = 12.8   | m = 5.63<br>s = 10.11 |
| Goal achievement         | m = 77.6<br>s = 43.19  | m = 78.3<br>s = 38.5 | m = 76.7<br>s = 38.6  | m = 66.3<br>s = 57.4 | m = 75<br>s = 29.66   | m = 71.7<br>s = 38.3  |
| Performance satisfaction | m = 84.0<br>s = 23.65  | m = 71.7<br>s = 35.3 | m = 79.3<br>m = 26.8  | m = 82.3<br>s = 30.6 | m = 70.6<br>s = 31.9  | m = 75<br>s = 29.7    |
| Mental effort            | m = 63.4<br>s = 19.1   | m = 13.0<br>s = 1054 | m = 44.5<br>s = 29.8  | m = 33<br>s = 15.72  | m = 32.6<br>s = 24.17 | m = 32.7<br>s = 20.1  |
| Physical effort          | m = 1.20<br>s = 2.68   | m = 1.33<br>s = 2.31 | m = 1.25<br>s = 2.38  | m = 0.0<br>s = 0.0   | m = 2.4<br>s = 4.83   | m = 1.5<br>s = 3.85   |
| Frustration level        | m = 15.2<br>s = 13.77  | m = 0.33<br>s = 0.58 | m = 9.63<br>s = 12.95 | m = 0.0<br>s = 0.0   | m = 14.0<br>s = 26.49 | m = 8.75<br>s = 21.3  |
| Stress level             | m = 3.8<br>s = 5.76    | m = 0.0<br>s = 0.0   | m = 2.38<br>s = 4.78  | m = 0.33<br>s = 0.58 | m = 10.8<br>s = 23.59 | m = 6.88<br>s = 18.64 |
| Overall Rating           | m = 35.3<br>s = 12.5   | m = 22.0<br>s = 7.56 | m = 30.3<br>s = 12.37 | m = 26.3<br>s = 7.79 | m = 29.3<br>s = 6.41  | m = 28.2<br>s = 6.57  |

A comparison of Task A versus Task B indicates had higher physical demand, time required, time pressure, goal achievement, performance satisfaction, mental effort, frustration level, and overall ratings for Task A. Participants had higher necessary thinking, task difficulty, physical effort and stress level ratings for Task B. Since the overall rating was higher for Task A, these results are consistent with the hypothesis for the enhanced interface in that it should reduce the perceived workload. Table 7.25 indicates that all of these results

are insignificant although performance satisfaction was close to significant ( $z = -1.82$ ,  $p = 0.06$ ).

Table 7.25. Tasks A versus Task B NASA-TLX

| Sub-Rating               | Statistic                 |
|--------------------------|---------------------------|
| Necessary thinking       | $z = -1.21$ , $p = 0.22$  |
| Task difficulty          | $z = -0.404$ , $p = 0.68$ |
| Physical demand          | $z = -0.40$ , $p = 1.0$   |
| Time required            | $z = -1.76$ , $p = 0.07$  |
| Time pressure            | $z = -0.53$ , $p = 0.59$  |
| Goal achievement         | $z = -0.73$ , $p = 0.46$  |
| Performance satisfaction | $z = -1.82$ , $p = 0.06$  |
| Mental effort            | $z = 0.0$ , $p = 1.0$     |
| Physical effort          | $z = -1.0$ , $p = 0.31$   |
| Frustration level        | $z = -0.36$ , $p = 0.71$  |
| Stress level             | $z = -.04$ , $p = 0.65$   |
| Overall Rating           | $z = -0.40$ , $p = 0.68$  |

As shown by Table 7.26, participants who performed Task E before Task F had higher ratings for necessary thinking, task difficulty, physical demand, time required, mental effort, frustration level and overall resources for Task E. Participants who performed Task F before Task E had higher time pressure, goal achievement, performance satisfaction, physical effort and stress level for Task E. Participants who performed Task E before Task F found higher ratings for necessary thinking, time required, goal achievement, performance satisfaction, mental effort, physical effort, and overall resources for Task F. Participants who performed Task F before Task E had higher ratings for task difficulty, physical demand,

time pressure, frustration level, and stress level for Task F. Appendix I Table 23 indicates that statistical analysis of these results was insignificant.

Table 7.26. Task E and Task F NASA-TLX (mean (m), standard deviations (s))

|                             | Task E                     |                            |                      | Task F                     |                            |                      |
|-----------------------------|----------------------------|----------------------------|----------------------|----------------------------|----------------------------|----------------------|
| Sub-Rating                  | Task E<br>before<br>Task F | Task F<br>before<br>Task E | Total                | Task E<br>before<br>Task F | Task F<br>before<br>Task E | Total                |
| Necessary<br>thinking       | m = 55.6<br>s = 26.1       | m = 33.4<br>s = 27.1       | m = 44.5<br>s = 27.7 | m = 42<br>s = 21.1         | m = 40.2<br>s = 26.1       | m = 41.1<br>s = 22.4 |
| Task<br>difficulty          | m = 51.2<br>s = 27.9       | m = 24.2<br>s = 26.4       | m = 37.7<br>s = 29.3 | m = 31.2<br>s = 24.0       | m = 38.6<br>s = 18.6       | m = 34.9<br>s = 20.6 |
| Physical<br>demand          | m = 1.8<br>s = 10          | m = 1.2<br>s = 1.3         | m = 1.5<br>s = 2.0   | m = 0.6<br>s = 0.9         | m = 2.2<br>s = 1.8         | m = 1.4<br>s = 1.6   |
| Time<br>required            | m = 36.4<br>s = 9.6        | m = 28<br>s = 29.9         | m = 32.2<br>s = 21.4 | m = 26.0<br>s = 18.7       | m = 14.8<br>s = 7.2        | m = 20.4<br>s = 14.6 |
| Time<br>pressure            | m = 6.4<br>s = 7.0         | m = 19.8<br>s = 34.9       | m = 13.1<br>s = 24.8 | m = 7.0<br>s = 8.8         | m = 10.8<br>s = 14.2       | m = 8.9<br>s = 11.3  |
| Goal<br>achievement         | m = 66.8<br>s = 39.1       | m = 75.8<br>s = 8.7        | m = 71.3<br>s = 27.1 | m = 63.8<br>s = 39.8       | m = 59.0<br>s = 25.0       | m = 61.4<br>s = 31.4 |
| Performance<br>satisfaction | m = 67.2<br>s = 26.4       | m = 71.6<br>s = 17.4       | m = 69.4<br>s = 21.2 | m = 63.0<br>s = 28.3       | m = 54.6<br>s = 21.1       | m = 58.8<br>s = 23.9 |
| Mental effort               | m = 49<br>s = 25.8         | m = 48.2<br>s = 31.8       | m = 48.6<br>s = 27.3 | m = 38.0<br>s = 18.5       | m = 36.4<br>s = 27.5       | m = 37.2<br>s = 22.1 |
| Physical<br>effort          | m = 2.0<br>s = 3.9         | m = 13.2<br>s = 22.3       | m = 7.6<br>s = 16.2  | m = 9.0<br>s = 16.0        | m = 2.4<br>s = 3.6         | m = 5.7<br>s = 11.5  |
| Frustration<br>level        | m = 20.2<br>s = 22.6       | m = 14.4<br>s = 20.1       | m = 17.3<br>s = 20.4 | m = 32.4<br>s = 46.0       | m = 34.8<br>s = 27.0       | m = 33.6<br>s = 35.6 |

Table 7.26, continued

|                |                      |                      |                      |                      |                      |                      |
|----------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Stress level   | m = 11.8<br>s = 25.3 | m = 14.2<br>s = 23.0 | m = 13.0<br>s = 22.8 | m = 12.8<br>s = 21.5 | m = 22.2<br>s = 25.7 | m = 17.5<br>s = 22.9 |
| Overall Rating | m = 33.5<br>s = 13.8 | m = 31.3<br>s = 13.6 | m = 32.4<br>s = 13.0 | m = 29.6<br>s = 13.1 | m = 28.7<br>s = 4.7  | m = 29.2<br>s = 9.3  |

A comparison of Task E versus Task F shows that Task E had higher ratings for necessary thinking, task difficulty, physical effort, time required, time pressure, goal achievement, performance satisfaction, mental effort, physical effort and the overall ratings. Task F had higher frustration level and stress level ratings. The reduction in the overall perceived workload from Task E to F could be attributed to the addition of the SES on the enhanced interface. The higher frustration levels for Task F could also be attributed to the addition of the SES. Table 7.27 indicates that none of these results was insignificant.

Table 7.27. Task E versus Task F NASA-TLX

| Sub-Rating               | Statistic             |
|--------------------------|-----------------------|
| Necessary thinking       | $z = -0.66, p = 0.50$ |
| Task difficulty          | $z = -0.35, p = 0.72$ |
| Physical demand          | $z = -0.10, p = 0.91$ |
| Time required            | $z = -1.63, p = 0.10$ |
| Time pressure            | $z = -0.88, p = 0.37$ |
| Goal achievement         | $z = -1.54, p = 0.12$ |
| Performance satisfaction | $z = -1.68, p = 0.92$ |
| Mental effort            | $z = -0.83, p = 0.40$ |
| Physical effort          | $z = -0.52, p = 0.59$ |
| Frustration level        | $z = -0.83, p = 0.4$  |
| Stress level             | $z = -0.21, p = 0.83$ |
| Overall Rating           | $z = -1.17, p = 0.24$ |

Table 7.28 indicates that there is one significant result when comparing workload between Tasks A and E. The overall perceived workload was higher for Task E. This result could be due to the fact that Task E was a teleoperation task while in Task A the robot was stationary ( $z = -2.38$ ,  $p = 0.02$ ). There were two significant results between Tasks B and F. The necessary thinking was higher for Task F ( $z = -2.24$ ,  $p = 0.02$ ). This is not surprising considering that the robot was mobile during Task F. Additionally, the physical effort was higher for Task F ( $z = -2.04$ ,  $p = 0.041$ ).

Table 7.28. Task Comparison NASA-TLX

| Sub-Rating               | A versus E                | B versus F                |
|--------------------------|---------------------------|---------------------------|
| Necessary thinking       | $z = 0.0$ , $p = 1.0$     | $z = -2.24$ , $p = 0.02$  |
| Task difficulty          | $z = -0.56$ , $p = 0.57$  | $z = -.14$ , $p = 0.88$   |
| Physical demand          | $z = -0.677$ , $p = 0.49$ | $z = -2.04$ , $p = 0.041$ |
| Time required            | $z = -1.4$ , $p = 0.16$   | $z = -0.14$ , $p = 0.88$  |
| Time pressure            | $z = -0.10$ , $p = 0.91$  | $z = -0.42$ , $p = 0.674$ |
| Goal achievement         | $z = -0.67$ , $p = 0.49$  | $z = -0.33$ , $p = 0.72$  |
| Performance satisfaction | $z = -1.12$ , $p = 0.26$  | $z = -1.85$ , $p = 0.06$  |
| Mental effort            | $z = -0.56$ , $p = 0.58$  | $z = 0.0$ , $p = 1.0$     |
| Physical effort          | $z = -1.21$ , $p = 0.22$  | $z = -0.36$ , $p = 0.71$  |
| Frustration level        | $z = -1.52$ , $p = 0.13$  | $z = -1.36$ , $p = 0.17$  |
| Stress level             | $z = -1.48$ , $p = 0.14$  | $z = -0.94$ , $p = 0.34$  |
| Overall Rating           | $z = -2.38$ , $p = 0.02$  | $z = -1.12$ , $p = 0.26$  |

Appendix I Table 24 provides the correlation analysis between the NASA-TLX ratings and the number of camera clicks. This analysis demonstrates that there is a positive correlation between the number of clicks and the perceived workload. Participants who

used the camera more experienced a higher perceived workload. For Task A, there was a positive correlation between the necessary thinking and the number of zoom-out clicks ( $r = 0.88$ ,  $p = 0.02$ ). Also for Task A, there was a positive correlation between the number of reset clicks and the time required ( $r = 0.893$ ,  $p = 0.02$ ) as well as a negative correlation with goal achievement ( $r = -0.83$ ,  $p = 0.04$ ). For Task B, there was a positive correlation between the number of zoom-in clicks and the time required ( $r = 0.861$ ,  $p = 0.013$ ) as well as the perceived mental effort ( $r = 0.975$ ,  $p = 0.0$ ). There was a negative correlation between the number of zoom-out clicks and goal achievement ( $r = -0.96$ ,  $p = 0.001$ ) as well as positive correlations between the number of pan clicks and physical effort ( $r = 0.77$ ,  $p = 0.04$ ), the number of tilt clicks and the frustration level ( $r = 0.788$ ,  $p = 0.035$ ), and the number of reset clicks and stress level ( $r = 0.959$ ,  $p = 0.001$ ) for Task B. A positive correlation existed for Task E between the number of zoom-out clicks and the time required ( $r = 0.664$ ,  $p = 0.036$ ) as well as with time pressure ( $r = 0.693$ ,  $p = 0.026$ ). A negative correlation was found between the number of zoom-in clicks and the goal achievement ( $r = -0.693$ ,  $p = 0.026$ ) for Task F. The remainder of the NASA-TLX ratings did not exhibit significant correlations with the number of camera clicks.

Appendix I Table 25 provides the correlation analysis between the NASA-TLX ratings and the number of map clicks. Negative correlations existed between the number of add icon clicks and the overall rating ( $r = -0.68$ ,  $p = 0.04$ ), the necessary thinking ( $r = -0.74$ ,  $p = 0.021$ ), and frustration level ( $r = -0.67$ ,  $p = 0.05$ ). For Task F, there was a positive correlation between the overall rating and the map clicks ( $r = 0.67$ ,  $p = 0.05$ ) as well as the add icon clicks ( $r = 0.691$ ,  $p = 0.039$ ). The negative correlations for Task E and the positive correlations for Task F suggest that there may not be a definite relationship between the NASA-TLX and the number of map clicks.

The correlation analysis between the NASA-TLX ratings and the number of SES clicks is provided in Appendix I Table 26. For Task B, there is a negative correlation between the necessary thinking and the scan clicks ( $r = -0.78$ ,  $p = 0.02$ ). A negative correlation between the necessary thinking and the pan right clicks ( $r = -0.636$ ,  $p = 0.04$ ), respectively. There is also a positive correlation between the task difficulty and the numbers of zoom-in clicks for Task B ( $r = 0.71$ ,  $p = 0.04$ ). Tasks B and F also demonstrate a positive correlation between mental effort and the scan clicks for Task B ( $r = 0.719$ ,  $p = 0.04$ ) and Task F ( $r = 0.66$ ,  $p = 0.04$ ). Therefore, the perceived effort actually increases with the use of the SES. For Task B, there was also a negative correlation between the number of zoom-out clicks and the frustration level ( $r = -0.72$ ,  $p = 0.04$ ) and the stress level ( $r = -0.72$ ,  $p = 0.05$ ).

The correlation analysis between the NASA-TLX and the task score are provided in Appendix I Table 27. There is a negative correlation between the driving directions score and the frustration level for Task A ( $r = -0.86$ ,  $p = 0.005$ ). For Task B, there is a negative correlation between the driving directions score and the time required ( $r = -0.87$ ,  $p = 0.01$ ), frustration level ( $r = -0.992$ ,  $p = 0.0$ ) and stress level ( $r = -1.0$ ,  $p = 0.0$ ). Additionally for Task B, there was a positive correlation between the robot placement score and the goal achievement ( $r = 0.75$ ,  $p = 0.03$ ). For Task E, there is a negative correlation between the overall score and the task difficulty ( $r = -0.64$ ,  $p = 0.05$ ), time required ( $r = -0.85$ ,  $p = 0.002$ ), time pressure ( $r = -0.75$ ,  $p = 0.012$ ), and frustration level ( $r = -0.83$ ,  $p = 0.002$ ). Additionally, there is a negative correlation between necessary thinking and cone color score ( $r = -0.74$ ,  $p = 0.01$ ) and a positive correlation between goal achievement and the cone placement score ( $r = 0.657$ ,  $p = 0.039$ ) for Task E. There was a positive correlation between the overall score and the mental effort for Task F ( $r = 0.66$ ,  $p = 0.04$ ). These negative correlations imply that the participants' perceived time demand, necessary thinking and frustration levels actually

reduce the task score. These positive correlations imply that the participants' perceived mental effort and goal achievement increase task score.

Appendix I Table 28 presents the correlation analysis between the NASA-TLX ratings and the task completion time. Only a single negative correlation existed between the driving direction time and frustration level for Task A ( $r = 0.808$ ,  $p = 0.015$ ). This isolated significant value suggests that it is a spurious result since no other ratings were significant.

Finally, Appendix I Table 29 presents an analysis of the MRQ values versus the NASA-TLX. Task A had two negative correlations, one between the short-term memory and the stress level ( $r = -0.846$ ,  $p = 0.008$ ) and one between the spatial emergent process and the task difficulty ( $r = -0.89$ ,  $p = 0.003$ ). There were several positive correlations between necessary thinking and the manual ( $r = 0.888$ ,  $p = 0.003$ ), spatial attentive ( $r = 0.796$ ,  $p = 0.018$ ), spatial categorical ( $r = 0.802$ ,  $p = 0.017$ ), and spatial positional ( $r = 0.754$ ,  $p = 0.031$ ) processes. Additionally, for Task A there was a negative correlation between the stress level and the visual lexical ( $r = -0.824$ ,  $p = 0.012$ ) and visual temporal ( $r = -0.796$ ,  $p = 0.018$ ) resources. There was a negative correlation between the spatial quantitative resources and time required ( $r = -0.89$ ,  $p = 0.003$ ), frustration level ( $r = -0.563$ ,  $p = 0.006$ ), and stress level ( $r = -0.89$ ,  $p = 0.003$ ) for Task B. There is a positive correlation between the overall perceived workload and short-term memory ( $r = 0.63$ ,  $p = 0.04$ ) and spatial attentive ( $r = 0.064$ ,  $p = 0.04$ ) demands for Task E. There were also positive correlations between the overall MRQ resources and the frustration level ( $r = 0.663$ ,  $p = 0.037$ ) and the stress level ( $r = 0.753$ ,  $p = 0.012$ ). For Task F, there was a negative correlation between the necessary thinking and the spatial positional process ( $r = -0.75$ ,  $p = 0.013$ ) and spatial quantitative process ( $r = -0.472$ ,  $p = 0.018$ ), and overall ratings ( $r = -0.75$ ,  $p = 0.013$ ). Additionally, there were negative correlations between the mental effort and spatial positional ( $r = -0.75$ ,  $p =$

0.013), spatial quantitative ( $r = -0.67$ ,  $p = 0.03$ ), and overall resources ( $r = -0.79$ ,  $p = 0.006$ ) for Task F. These results indicate that there are relationships between some of the MRQ categories and the NASA-TLX, such as the frustration, stress, spatial attentive, spatial quantitative, spatial positional, necessary thinking, and mental demand.

### **Spatial Reasoning**

The purpose of the spatial reasoning analysis was to determine if participants' level of spatial reasoning effected task performance. The results of the spatial reasoning test were discussed in the participant demographics section. Only the correlations between other variables and the spatial reasoning scores are presented here. The spatial reasoning test is located in Appendix B. No significant correlations existed between spatial reasoning and the number of camera or map clicks.

With respect to the correlation analysis between spatial reasoning and the number of SES clicks, there were two positive correlations for Task F. A positive correlation existed between the spatial reasoning ratings and the number of scan clicks ( $r = 0.683$ ,  $p = 0.037$ ) as well as the number of reset ( $r = 0.894$ ,  $p = 0.026$ ) clicks. This result means that participants with higher spatial reasoning used the camera more to complete the task. These results are found in Appendix I Table 30. The correlation analysis between the spatial reasoning score and task time had one significant result with the time to find the robot ( $r = 0.857$ ,  $p = 0.007$ ) for Task A. The correlation analysis performed between the participants' spatial reasoning score and task scores found no significant correlations.

Only one correlation existed between the spatial reasoning scores and the MRQ ratings. A negative correlation with the spatial emergent process ( $r = -0.791$ ,  $p = 0.006$ )

existed for Task F. Since this result was isolated, it was most likely spurious. No other comparisons were significant.

Finally, a correlation analysis was performed between the spatial reasoning and the NASA-TLX ratings. The overall results indicate that perceived workload decreases with higher spatial reasoning scores for Task F. These results can be found in Appendix I Table 31. There were no significant correlations for Tasks B and E. For Task A, there was a positive correlation with the frustration level ( $r = 0.81$ ,  $p = 0.014$ ). For Task F, there were negative correlations with necessary thinking ( $r = 0.67$ ,  $p = 0.03$ ), task difficulty ( $r = -0.67$ ,  $p = 0.031$ ), frustration level ( $r = -0.71$ ,  $p = 0.02$ ) and the overall perceived workload rating ( $r = -0.91$ ,  $p = 0.0$ ).

### **Post-Task Questionnaire**

The purpose of the post-task questionnaire was to obtain the participants' overall reaction to the task, such as the perceived time pressure, task difficulty, and system capabilities. The actual questionnaire is located in the Appendix E. As shown in Appendix E, the likert rating scale for the post-task questionnaire was 0 to 5.

Table 7.29 presents the mean and standard deviations for the task rating specific questions dependent upon task order. Participants who performed Task A before Task B rated clarity level, stimulation level, and frustration level higher for Task A. Participants who performed Task B before Task A rated the difficulty level higher for Task A. Participants who performed Task A before Task B rated clarity and stimulation levels higher for Task B. Participants who performed Task B before Task A rated the frustration level higher for Task B. There were equivalent ratings for the difficulty level for Task B. Appendix I Table 32 indicates that significant results existed based on task order existed for Task A stimulation (z

= -2.39,  $p = 0.016$ ) and frustration ( $z = -1.91$ ,  $p = 0.05$ ) levels. These results indicate that task stimulation and frustration levels were significantly higher when Task A was performed before Task B.

Table 7.29. Task A and Task B Task Specific Ratings (mean (m), standard deviations (s))

|                      | Task A                     |                            |                      | Task B                     |                            |                      |
|----------------------|----------------------------|----------------------------|----------------------|----------------------------|----------------------------|----------------------|
| Question             | Task A<br>before<br>Task B | Task B<br>before<br>Task A | Total                | Task A<br>before<br>Task B | Task B<br>before<br>Task A | Total                |
| difficulty<br>level  | m = 3.00<br>s = 0.82       | m = 2.75<br>s = 2.22       | m = 2.88<br>s = 1.55 | m = 3.33<br>s = 0.58       | m = 3.33<br>s = 1.21       | m = 3.33<br>s = 1.00 |
| clarity level        | m = 4.00<br>s = 1.15       | m = 2.00<br>s = 1.50       | m = 3.13<br>s = 1.55 | m = 4.33<br>s = 1.15       | m = 3.67<br>s = 1.03       | m = 3.89<br>s = 1.05 |
| stimulation<br>level | m = 4.0<br>s = 1.0         | m = 1.25<br>s = 1.89       | m = 2.75<br>s = 2.05 | m = 4.00<br>s = 0.0        | m = 2.72<br>s = 1.47       | m = 3.89<br>s = 1.17 |
| frustration<br>level | m = 3.0<br>s = 2.45        | m = 1.00<br>s = 2.00       | m = 2.00<br>s = 2.33 | m = 3.0<br>s = 2.65        | m = 3.67<br>s = 2.16       | m = 3.44<br>s = 2.19 |

A comparison of Task A versus Task B task specific ratings showed that Task B was rated higher for all of the ratings. Participants felt that Task B was more difficult, more frustrating but clearer and more stimulating than Task A. Table 7.30 indicates that none of these differences were significant.

Table 7.30. Task A versus Task B Task Specific Ratings (Statistics)

| Question          | Statistic             |
|-------------------|-----------------------|
| difficulty level  | $z = -0.27, p = 0.78$ |
| clarity level     | $z = -1.34, p = 0.18$ |
| stimulation level | $z = -1.41, p = 0.15$ |
| frustration level | $z = -1.28, p = 0.19$ |

Table 7.31 provides the mean and standard deviations for general task questions for Tasks A and B dependent upon task order. Participants who performed Task A before Task B rated their ability to understand data, correct errors, and the perceived level of control over the system higher for Task A. Participants who performed Task B before Task A rated their ability to complete tasks higher for Task A. Participants who performed Task A before Task B rated their ability to correct errors and control the system higher for Task B. Participants who performed Task B before Task A felt they better understood the data displays and were more confident in the ability to complete Task B. Appendix I Table 33 indicates that the only significant result existed for the ability to correct errors for Task A when Task A was completed first ( $z = -1.91, p = 0.05$ ).

Table 7.31. Task A and Task B General Questions (mean (m), standard deviations (s))

| Question                      | Task A               |                      |                      | Task B               |                      |                      |
|-------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                               | Task A before Task B | Task B before Task A | Total                | Task A before Task B | Task B before Task A | Total                |
| understand and interpret data | m = 4.00<br>s = 1.00 | m = 3.25<br>s = 2.22 | m = 3.75<br>s = 1.58 | m = 4.0<br>s = 0.0   | m = 4.5<br>s = 0.55  | m = 4.33<br>s = 0.50 |

Table 7.31, continued

|                               |                      |                      |                      |                      |                      |                      |
|-------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| error correction capabilities | m = 4.50<br>s = 0.58 | m = 3.00<br>s = 2.00 | m = 3.50<br>s = 1.69 | m = 4.33<br>s = 0.58 | m = 4.0<br>s = 1.09  | m = 4.11<br>s = 0.93 |
| system control                | m = 2.25<br>s = 2.63 | m = 3.00<br>s = 2.00 | m = 2.38<br>s = 2.13 | m = 4.33<br>s = 0.58 | m = 4.0<br>s = 1.09  | m = 4.11<br>s = 0.93 |
| ability to complete tasks     | m = 4.00<br>s = 1.00 | m = 4.50<br>s = 0.58 | m = 4.38<br>s = 0.52 | m = 4.67<br>s = 0.58 | m = 4.83<br>s = 0.41 | m = 4.78<br>s = 0.44 |

A comparison of Task A versus Task B indicates from Table 7.31 that all questions were rated higher for Task B. Participants better understood the data, could correct their errors, better control the system and complete Task B over Task A. This may indicate that the enhanced interface was preferred by participants. Table 7.32 demonstrates that none of these differences was significant.

Table 7.32. Task A versus Task B General Questions (Statistics)

| Question                      | Statistic               |
|-------------------------------|-------------------------|
| understand and interpret data | $z = -0.272, p = 0.785$ |
| error correction capabilities | $z = -1.345, p = 0.180$ |
| system control                | $z = -1.414, p = 0.157$ |
| ability to complete tasks     | $z = -1.289, p = 0.197$ |

Table 7.33 provides the mean and standard deviations for the ratings of the system capabilities. The table indicates that participants who performed Task A before Task B rated the system as more flexible for Task A. They also agreed that the system provided the ability to easily complete the task. Participants who performed Task B before Task A rated the system higher for system speed, power level, time constraints and support information

satisfaction for Task A. Also from Table 7.33, participants who performed Task A before Task B rated the system higher for system speed, power level, rigidity, ease of task completion, and support information satisfaction for Task B. Participants who performed Task B before Task A felt that the time constraints were higher. Appendix I Table 34 indicates that none of these results based upon task order were significant.

Table 7.33. Task A and Task B System Capability Ratings (mean (m), standard deviations (s))

|                                  | Task A                     |                            |                      | Task B                     |                            |                      |
|----------------------------------|----------------------------|----------------------------|----------------------|----------------------------|----------------------------|----------------------|
| Question                         | Task A<br>before<br>Task B | Task B<br>before<br>Task A | Total                | Task A<br>before<br>Task B | Task B<br>before<br>Task A | Total                |
| system<br>speed                  | m = 1.75<br>s = 2.36       | m = 3.75<br>s = 1.26       | m = 2.75<br>s = 2.05 | m = 3.00<br>s = 2.65       | m = 2.67<br>s = 1.96       | m = 2.78<br>s = 2.05 |
| power<br>level                   | m = 2.25<br>s = 2.63       | m = 3.75<br>s = 2.50       | m = 3.00<br>s = 2.51 | m = 3.00<br>s = 2.65       | m = 1.83<br>s = 2.23       | m = 2.22<br>s = 2.28 |
| rigidity                         | m = 2.00<br>s = 2.63       | m = 1.00<br>s = 2.00       | m = 1.63<br>s = 2.26 | m = 4.33<br>s = 0.58       | m = 2.0<br>s = 2.1         | m = 2.78<br>s = 2.05 |
| ease of<br>task<br>completion    | m = 4.25<br>s = 0.50       | m = 3.75<br>s = 1.26       | m = 4.00<br>s = 0.93 | m = 4.33<br>s = 0.58       | m = 3.5<br>s = 2.1         | m = 3.78<br>s = 1.72 |
| time<br>constraints              | m = 3.00<br>s = 2.00       | m = 2.75<br>s = 2.22       | m = 2.88<br>s = 1.96 | m = 2.67<br>s = 2.31       | m = 3.17<br>s = 1.72       | m = 3.00<br>s = 1.80 |
| support<br>info.<br>satisfaction | m = 4.50<br>s = 0.58       | m = 4.50<br>s = 0.58       | m = 4.50<br>s = 0.53 | m = 4.33<br>s = 0.58       | m = 3.33<br>s = 1.96       | m = 3.67<br>s = 1.66 |

Participants felt that Task A was easier to complete, more powerful and were more satisfied with support information, while they felt the system speed, flexibility and time to complete the task were better for Task B. Table 7.34 indicates that none of these differences was significant.

Table 7.34. Task A versus Task B System Capability Ratings (SPSS)

| Question                         | Statistic             |
|----------------------------------|-----------------------|
| system speed                     | $z = -0.55, p = 0.58$ |
| power level                      | $z = -1.34, p = 0.18$ |
| rigidity                         | $z = -1.60, p = 0.10$ |
| ease of task completion          | $z = -1.41, p = 0.15$ |
| time constraints                 | $z = -1.34, p = 0.18$ |
| support information satisfaction | $z = -1.6, p = 0.11$  |

Table 7.35 provides a comparison of the responses based on Tasks E and F order for task specific questions. Participants who performed Task F before Task E rated difficulty level, clarity level, stimulation level, and frustration level higher for Task E. Participants who performed Task E before Task F rated clarity, stimulation, and frustration levels higher for Task F. The ratings for the difficulty level for Task F were equivalent independent of task presentation order. This shows that typically the second run of a task received higher ratings. Appendix I Table 35 indicates that none of these differences were significant.

Table 7.35. Task E and Task F Task Specific Ratings (mean (m), standard deviations (s))

|                     | Task E                     |                            |                      | Task F                     |                            |                      |
|---------------------|----------------------------|----------------------------|----------------------|----------------------------|----------------------------|----------------------|
| Question            | Task E<br>before<br>Task F | Task F<br>before<br>Task E | Total                | Task E<br>before<br>Task F | Task F<br>before<br>Task E | Total                |
| difficulty<br>level | m = 2.0<br>s = 1.22        | m = 2.8<br>s = 1.10        | m = 2.40<br>s = 1.17 | m = 3.2<br>s = 0.84        | m = 3.2<br>s = 0.84        | m = 3.20<br>s = 0.79 |
| clarity level       | m = 2.2<br>s = 1.30        | m = 3.40<br>s = 0.89       | m = 2.80<br>s = 1.23 | m = 3.40<br>s = 0.89       | m = 3<br>s = 0.00          | m = 3.20<br>s = 0.63 |

Table 7.35, continued

|                   |                     |                      |                      |                      |                     |                      |
|-------------------|---------------------|----------------------|----------------------|----------------------|---------------------|----------------------|
| stimulation level | m = 3.2<br>s = 1.78 | m = 3.60<br>s = 2.07 | m = 3.40<br>s = 1.84 | m = 4.40<br>s = 0.55 | m = 2.6<br>s = 2.41 | m = 3.50<br>s = 1.90 |
| frustration level | m = 2.8<br>s = 1.78 | m = 3.00<br>s = 2.00 | m = 2.90<br>s = 1.79 | m = 3.40<br>s = 1.34 | m = 1.8<br>s = 2.49 | m = 2.60<br>s = 2.07 |

Participants felt that Task E was more frustrating although Task F was easier, more clear and stimulating. Table 7.36 indicates that none of these results were significant.

Table 7.36. Task E versus Task F Task Specific Ratings (SPSS)

| Question          | Statistic               |
|-------------------|-------------------------|
| difficulty level  | $z = -1.807, p = 0.071$ |
| clarity level     | $z = -1.08, p = 0.276$  |
| stimulation level | $z = -0.272, p = 0.785$ |
| frustration level | $z = -0.552, p = 0.58$  |

Table 7.37 is a comparison of the responses to Tasks E and F general ratings dependent upon task presentation order. Participants who performed Task E before Task F rated their understanding and interpretation of data higher for Task E. Participants who performed Task F before Task E rated their error correction capabilities, system control and ability to complete tasks higher for Task E. Participants who performed Task E before Task F rated their understanding and interpretation of data and error correction capabilities higher for Task F. Participants who performed Task F before Task E rated the system control and ability to complete tasks higher for Task F. Appendix I Table 36 indicates that none of these differences were significant.

Table 7.37. Task E and Task F General Ratings (mean (m), standard deviations (s))

| Question                      | Task E               |                      |                      | Task F               |                      |                      |
|-------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                               | Task E before Task F | Task F before Task E | Total                | Task E before Task F | Task F before Task E | Total                |
| understand and interpret data | m = 3.4<br>s = 1.94  | m = 2.60<br>s = 2.41 | m = 3.00<br>s = 2.11 | m = 4.20<br>s = 0.45 | m = 3.4<br>s = 1.95  | m = 3.80<br>s = 1.40 |
| error correction capabilities | m = 3.2<br>s = 1.78  | m = 4.20<br>s = 0.45 | m = 3.70<br>s = 1.34 | m = 4.20<br>s = 0.45 | m = 3.4<br>s = 1.95  | m = 3.80<br>s = 1.40 |
| system control                | m = 2.8<br>s = 1.78  | m = 3.40<br>s = 1.95 | m = 3.10<br>s = 1.79 | m = 3.40<br>s = 1.34 | m = 4.2<br>s = 0.45  | m = 3.80<br>s = 1.03 |
| ability to complete tasks     | m = 4.2<br>s = 0.44  | m = 4.60<br>s = 0.55 | m = 4.40<br>s = 0.52 | m = 4.20<br>s = 0.45 | m = 4.8<br>s = 0.45  | m = 4.50<br>s = 0.53 |

Participants felt that the enhanced interface was easier to control and use to complete tasks during Task F. Participants also felt that it was easier to understand and interpret data as well as correct errors during Task F. Table 7.38 indicates that none of these results was significant.

Table 7.38. Task E versus Task F General Ratings (SPSS)

| Question                      | Statistic              |
|-------------------------------|------------------------|
| understand and interpret data | $z = -1.069, p = 0.28$ |
| error correction capabilities | $z = -0.272, p = 0.78$ |
| system control                | $z = -1.604, p = 0.10$ |
| ability to complete tasks     | $z = -0.577, p = 0.56$ |

Table 7.39 is a comparison of the responses to Tasks E and F system capability ratings dependent upon task presentation order. Participants who performed Task E before Task F rated the system more rigid for Task E. All other questions were rated higher by participants who performed Task F before Task E. Participants who performed Task E before Task F rated Task F higher for power level, rigidity, ease of task completion and support information satisfaction for Task F. Participants who performed task F before Task E rated the system speed and time constraints higher for Task F. Appendix I Table 37 indicates that none of these differences was statistically significant.

Table 7.39. Task E and Task F System Capability Ratings (mean (m), standard deviations (s))

| Question                   | Task E               |                      |                      | Task F               |                      |                      |
|----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                            | Task E before Task F | Task F before Task E | Total                | Task E before Task F | Task F before Task E | Total                |
| system speed               | m = 2.6<br>s = 1.94  | m = 3.20<br>s = 1.10 | m = 2.90<br>s = 1.52 | m = 1.40<br>s = 2.19 | m = 2.2<br>s = 2.28  | m = 1.80<br>s = 2.15 |
| power level                | m = 1.8<br>s = 2.48  | m = 2.40<br>s = 2.51 | m = 2.10<br>s = 2.38 | m = 3.00<br>s = 2.00 | m = 2.8<br>s = 2.59  | m = 2.90<br>s = 2.18 |
| rigidity                   | m = 3<br>s = 2       | m = 1.40<br>s = 2.19 | m = 2.20<br>s = 2.15 | m = 2.60<br>s = 1.95 | m = 1.6<br>s = 1.67  | m = 2.10<br>s = 1.79 |
| ease of task completion    | m = 3.6<br>s = 0.89  | m = 4.20<br>s = 0.45 | m = 3.90<br>s = 0.74 | m = 3.80<br>s = 1.10 | m = 3.4<br>s = 1.95  | m = 3.60<br>s = 1.51 |
| time constraints           | m = 2.4<br>s = 1.67  | m = 3.80<br>s = 1.10 | m = 3.10<br>s = 1.52 | m = 2.80<br>s = 1.79 | m = 3.6<br>s = 1.52  | m = 3.20<br>s = 1.62 |
| support info. satisfaction | m = 4<br>s = 1.22    | m = 4.20<br>s = 0.45 | m = 4.10<br>s = 0.88 | m = 4.00<br>s = 1.22 | m = 3.8<br>s = 1.10  | m = 3.90<br>s = 1.10 |

Participants felt Task E had better flexibility, speed, ease of task completion and support information than Task F. While, participants felt better about the power level and satisfaction with the time to complete the task for Task F. Table 7.40 demonstrates that these results were not significant.

Table 7.40. Task E versus Task F System Capability Ratings (SPSS)

| Question                         | Statistic               |
|----------------------------------|-------------------------|
| system speed                     | $z = -1.769, p = 0.077$ |
| power level                      | $z = -1.63, p = 0.10$   |
| rigidity                         | $z = -0.272, p = 0.78$  |
| ease of task completion          | $z = -0.378, p = 0.705$ |
| time constraints                 | $z = -0.272, p = 0.705$ |
| support information satisfaction | $z = -1, p = 0.317$     |

An analysis was computed between tasks for the same interface. As Table 7.41 indicates that no significant differences existed between the post-task ratings for Task A versus Task E nor Task B versus Task F.

Table 7.41. Task Comparison Statistics

| Question                      | A versus E             | B versus F             |
|-------------------------------|------------------------|------------------------|
| difficulty level              | $z = -1.105, p = 0.27$ | $z = -0.557, p = 0.58$ |
| clarity level                 | $z = -0.27, p = 0.79$  | $z = -1.41, p = 0.16$  |
| stimulation level             | $z = -0.55, p = 0.58$  | $z = -0.73, p = 0.47$  |
| frustration level             | $z = -0.91, p = 0.36$  | $z = -0.408, p = 0.68$ |
| understand and interpret data | $z = -0.7, p = 0.46$   | $z = -1, p = 0.32$     |
| error correction capabilities | $z = -0.18, p = 0.85$  | $z = -0.44, p = 0.66$  |

|                                  |                        |                          |
|----------------------------------|------------------------|--------------------------|
| system control                   | $z = -0.55, p = 0.58$  | $z = -0.44, p = 1.00$    |
| ability to complete tasks        | $z = 0.0, p = 1.0$     | $z = -0.57, p = 0.56$    |
| system speed                     | $z = -0.184, p = 0.85$ | $z = 0.00, p = 1.000$    |
| power level                      | $z = -0.13, p = 0.89$  | $a = -90.921, p = 0.357$ |
| rigidity                         | $z = -1.841, p = 0.07$ | $z = -0.75, p = 0.45$    |
| ease of task completion          | $z = -0.55, p = 0.58$  | $z = -0.44, p = 0.66$    |
| time constraints                 | $z = -0.141, p = 0.89$ | $z = -1.633, p = 0.102$  |
| support information satisfaction | $z = -0.81, p = 0.41$  | $z = -1.0, p = 0.317$    |

### Pre- versus Post-Experiment Display Comparison

Each participant completed a pre-experiment questionnaire that included several potential interface display screens. The participants were asked to select a preferred display view from the choices both prior to the study and upon study completion. Figure 7.1 provides screen shots of the displays that the participants viewed before and after the experiment.

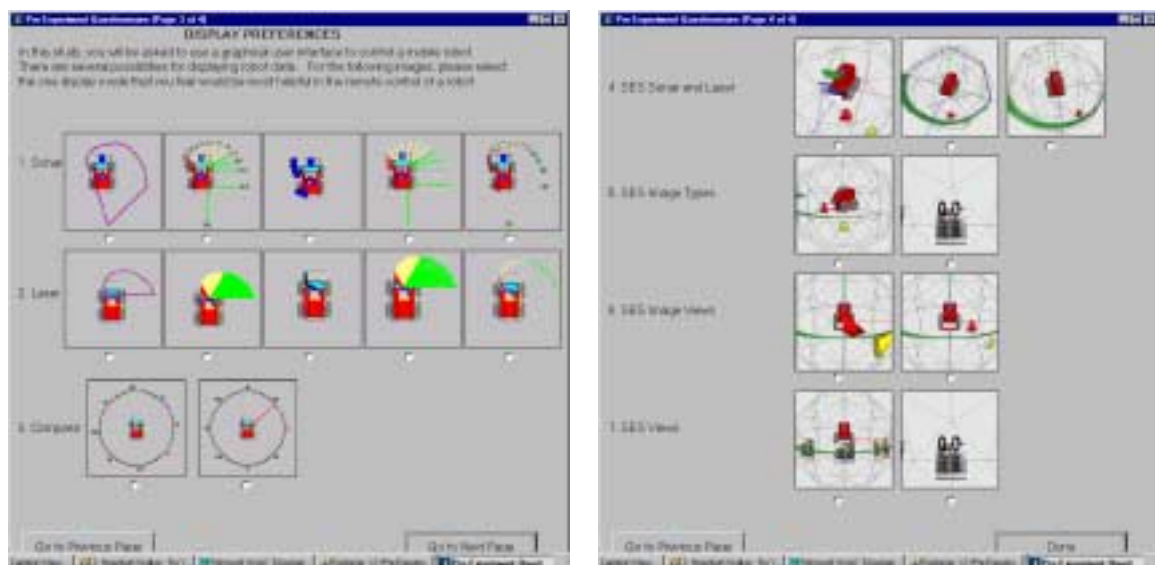


Figure 7.1. Display View Electronic Survey

The preference survey results are provided in Table 7.42. It was found that the majority of participants chose the default view for the compass, SES sonar, SES image types, SES image views, and SES views before the study began. After the study completion, the majority of the participants selected the default view for all the displays they used during the study. Table 7.42 table indicates the display preferences before and after the study as well as the default view.

Table 7.42. Pre- versus Post-Experiment Comparison

|                                          | Pre-Experimental                                     | Post-Experimental                      |
|------------------------------------------|------------------------------------------------------|----------------------------------------|
| P1 (Sonar)<br>Default – view 2           | 3 - view 2<br>5 - view 5<br>2 - no preference        | 9 – view 2<br>1 – view 5               |
| P2 (Laser)<br>Default – view 2           | 2 – view 1<br>2 – view 2<br>3 - view 4<br>3 – view 5 | 5 - view 2<br>5 – view 4               |
| P3 (Compass)<br>Default – view 2         | 1 - view 1<br>9 - view 2                             | 10 - view 2                            |
| P4 (SES Sonar)<br>Default – view 2       | 6 – view 2<br>4 - view 3                             | 1 – view 1<br>7 – view 2<br>2 - view 3 |
| P5 (SES Image Types)<br>Default – view 1 | 9 - view 1<br>1 - view 2                             | 10 - view 1                            |
| P6 (SES Image Views)<br>Default – view 2 | 1 - view 1<br>9 – view 2                             | 10 - view 2                            |
| P7 (SES Views)<br>Default - view 1       | 9 - view 1<br>1 - view 2                             | 10 - view 1                            |

### **Post-Experiment Questionnaire**

The purpose of this analysis was to determine the rating of the overall user interface across several categories. The actual questionnaire is located Appendix F. As the Appendix indicates, the rating scale was 0 to 5.

#### ***Question Set One***

The first question set assessed the participant's overall system reaction such as frustration level, power, stimulation, ease of use, and flexibility amount. The overall reaction to the system was slightly wonderful ( $m = 4, s = 0.82$ ). The system was easy to use ( $m = 3.5, s = 0.97$ ). Participants were satisfied with the system ( $m = 3.5, s = 1.08$ ) and rated the overall system power as neutral ( $m = 3.1, s = 1.2$ ). Finally, participants found the overall system slightly stimulating ( $m = 3.9, s = 0.88$ ) and flexible ( $m = 3.5, s = 1.08$ ).

#### ***Question Set Two***

The second question set dealt with issues related to the displayed screen information. The results show that participants rated reading characters as slightly easy ( $m = 4.4, s = 0.7$ ). The information organization was rated as slightly clear ( $m = 3.8, s = 1.0$ ) while the display layout was rated as slightly helpful ( $m = 4.1, s = 0.9$ ). The buttons were considered slightly easy to use ( $m = 4.0, s = 1.3$ ) and there were rarely clearly marked exits from the system ( $m = 2.5, s = 2.2$ ). The interface almost always provided a simple and natural dialogue ( $m = 4.3, s = 0.7$ ). The participants found the controls and actions to be somewhat compatible ( $m = 3.4, s = 0.5$ ). The participants felt that it was somewhat easy to locate the necessary task – specific information ( $m = 3.9, s = 0.9$ ). They also felt that the display layout neither

simplified nor complicated the task ( $m = 3.1, s = 1.29$ ). The participants felt that the display clutter was neutral ( $m = 3.1, s = 0.99$ ).

The participants felt items paired for direct comparison were almost slightly clear ( $m = 3.6, s = 1.5$ ). They also felt the indication of the active window was slightly clear ( $m = 3.9, s = 1.99$ ). The display contrast with the background was rated as slightly excellent ( $m = 3.9, s = 1.29$ ). The cursor was clearly distinctive ( $m = 4.2, s = 1.03$ ), and the display elements were slightly clear ( $m = 4.2, s = 1.03$ ). The group demarcation ( $m = 3.6, s = 1.51$ ) was slightly clear as well as the menu distinctions ( $m = 4, s = 1.49$ ) and screen sequencing ( $m = 3.9, s = 1.6$ ).

### ***Question Set Three***

The third question set probed system features such as terminology, message position, help and documentation as well as error messages. Participants felt the use of terms throughout the system was slightly consistent ( $m = 4.1, s = 0.57$ ). The position of messages on the screen was found to be somewhat slightly consistent ( $m = 3.5, s = 1.43$ ). The input prompts were slightly clear ( $m = 3.7, s = 1.6$ ). The computer almost always informs the user about its progress ( $m = 4, s = 0.94$ ). The help and documentation were rated as neutral ( $m = 3.3, s = 1.9$ ). The error messages were rated as slightly unhelpful ( $m = 1.9, s = 2.1$ ).

### ***Question Set Four***

Questions set four inquired regarding the ability to learn the system such as the time, getting started, system functionality, and steps to complete a task. Learning to operate the system was fairly easy ( $m = 4.3, s = 0.67$ ). Participants rated getting started with the system as slightly easy ( $m = 4, s = 1.15$ ). The system learning time was somewhat fast ( $m = 3.8, s =$

1.14). Remembering rules about entering commands was slightly easy ( $m = 4.4$ ,  $s = 0.7$ ). The thoughts about system functionality and capability were somewhat above neutral ( $m = 3.4$ ,  $s = 1.17$ ). Participants felt that they could almost always perform tasks in a straightforward manner ( $m = 3.6$ ,  $s = 0.7$ ). Participants were neutral regarding their opinion of the number of steps required per task ( $m = 3$ ,  $s = 1.05$ ). Participants thought that the keyboard shortcuts were not very helpful ( $m = 1.3$ ,  $s = 1.42$ ) and that the steps to complete a task almost always followed a logical sequence ( $m = 4$ ,  $s = 0.82$ ).

### ***Question Set Five***

Question set five explored the participants' reaction to the system capabilities. Questions involved qualities such as system speed, rate at which information is displayed and correcting mistakes. The system speed ( $m = 2.5$ ,  $s = 1.08$ ) and response time ( $m = 2.6$ ,  $s = 1.17$ ) received ratings leaning towards slightly slow. Participants were neutral regarding the information display rate ( $m = 3.1$ ,  $s = 0.99$ ) and the system failures ( $m = 3$ ,  $s = 1.56$ ). Participants felt that novices could accomplish tasks after proper training with slight ease ( $m = 3.7$ ,  $s = 1.34$ ). They also felt it was slightly easy to correct mistakes ( $m = 3.7$ ,  $s = 1.34$ ).

### ***Question Set Six***

Questions set six dealt with user comments such as system pleasantness, system satisfaction and how comfortable the participant was using the system. Participants agreed that they liked using the system ( $m = 4.1$ ,  $s = 0.88$ ). They also agreed that the system interface was pleasant ( $m = 4.1$ ,  $s = 0.88$ ) and overall were satisfied with the system ( $m = 4.1$ ,  $s = 1.1$ ). Participants strongly agreed that it was easy to learn this system ( $m = 4.5$ ,  $s =$

0.53). Additionally, participants agreed that they felt comfortable using this system ( $m = 4.4$ ,  $s = 0.7$ ).

### **Qualitative Results/ User Comments**

Each participant completed post task and post experiment questionnaires and were given the opportunity to provide comments. The post-experiment questionnaire, contained very specific questions such as, "list three negative aspects of using this system", "list three positive aspects of using this system", "what was your most common mistake", "what part of the evaluation was particularly frustrating or irritating", and "how would you suggest improving the interface". The results of the participants responses to these questions and comments are provided here.

### **Training**

The initial training for the interface evaluation included a PowerPoint presentation with screen shots and videos of the interface and robot. Each display element was thoroughly explained and demonstrated. The displays highlighted included sonar, laser, compass, camera, SES, remote drive, as well as the move to point and move to object behaviors.

Overall, the reaction to the training was very positive and most participants stated that it was concise, clear, and easy to understand. The one negative was that several participants felt a bit confused by the SES presentation. The following is a sample of some of the more prevalent comments.

- “The training made more sense after getting on the interface.”

- “I felt the training was a little slow and the SES would help get tasks done more quickly.”
- “I felt the directions to use the interface were straightforward and the visuals were excellent, but I suppose the SES was not clearly explained to me because I did not use it properly.”

### **Task Reactions**

As previously stated, each participant was asked to perform four tasks. The participants were provided written instructions as well as a verbally task explanation. The overwhelming response was that the task descriptions were concise, clear, and easy to understand. There were several statements that the task was reasonable, although a couple of participants felt it was challenging and more of a mental test. Other statements indicated the interface was easy to learn and became easier with time. The following is a sampling of participant reactions to the tasks.

- “As an inexperienced robotics user, I felt comfortable with the interface and the instructions on how to operate the interface.”
- “The instructions for completing the tasks during the initial study visit were very clear and concise. Having two chances to complete tasks was helpful, because the second time was definitely much easier and I had much more confidence. The tools used to assess landmarks, obstacles, and destinations were very easy even for someone like myself who wasn't quite sure what "sonar" meant on the first day. Driving the robot was just plain fun! And the researcher did a great job of making me feel comfortable and accomplished.”
- “I felt a sense of accomplishment after we had reached our goal.”

- “It appeared that the second task was a bit easier than the first because I learned how to navigate the robot better. After correcting the errors on the first task, everything made more sense so after moderate training would be an easy task to do.”

### **Camera Display**

The camera display was one of the most frequently used display elements during the interface evaluation tasks. As previously mentioned, although the robot possessed two cameras only the forward facing device was functional. Due to this issue, there was an overwhelming number of requests for the camera to view backwards or a full 360 degrees. The interface and the robot communicated over a wireless network and because of this there were several statements regarding the slow feedback. Many participants stated that at times they relied heavily on the camera to make decisions about the task and this was sometimes to their detriment. One consequence of the ease of using the camera interface was that participants neglected to consult the other displays before making decisions. There was one excellent suggestion that would possibly improve the camera display. The suggestion was to add a compass reading to the viewing window in order to show the direction the robot was viewing. One participant actually thought the robot turned with the camera so this would also be resolved with the addition of the compass heading to the camera view. The vast majority of participants felt the camera was the most useful sensor device for completing all tasks. There were a couple of participants who encountered difficulty determining color while using the camera. Some of these issues were resolved when the SES was present. It was assumed that these participants could possibly be colorblind. One example of user feedback was,

“Increase utility of camera interface, the laser and sonar readings weren't as helpful. These other two systems provided slightly distorted or confusing feedback and were not as helpful as actual camera and SES system in locating objects.”

### **Sonar, Laser, and Compass Displays**

The reactions to the sonar, laser, and compass displays were overwhelmingly negative. The majority of the responses were based upon the large error found in the sensory feedback. Participants stated that at times the compass display *lied* to them about the robot's true heading. Participants also felt that the laser was more reliable than the sonar but because it was only in the front of the robot. The compass was the most used sensor frequently because it was the most familiar. Some participants also felt that the sonar and laser visual displays were difficult to use. There were divergent views regarding sensory preferences and some participants preferred the sonar over the laser and vice versa. A sampling of the user comments are provided:

- “I got confused when trying to consolidate information from the sonar and the laser.”
- “The color coded display for the sonar and laser was useful.”
- “I liked the use of both sonar, laser, and camera. The camera will pick up most things, but the stuff behind and to the sides can't always be seen...so using the sonar and laser, I could understand where things are.”
- “I didn't really like the radar (sonar, laser). It didn't irritate me but I felt it was not really necessary because the sonar gave more accurate data and the camera gave even more data of the conditions around the robot.”

## **Sensory EgoSphere Display**

The primary difference between the original user interface and the enhanced user interface was the addition of the SES. The participants' reactions to the SES were almost exactly split down the middle. One overwhelming response was that the participants were confused by the SES and therefore preferred to use the camera to locate objects. Participants who preferred the SES liked all presentations of the sensory data in a compact form in one display. One negative comment was that at times the SES found additional icons because of image processing errors, therefore participants felt that it could not be trusted. Some of the more significant observations were:

- “The SES was not too helpful. I relied more on the camera than anything else. At the end of the first task, the SES said that there was an orange cone, I saw a green cone, but actually it was a yellow cone. So the SES was wrong in relation to the color of the cone. Also, it was hard to see the distances in relation to the robot according to the SES.”
- “The SES was a little helpful but not significantly.”
- “The superimposition on the SES helped to determine the line to objects, it was nice but not necessary.”
- “The SES missing from some tasks was a negative and it would help to locate objects and the robot.”
- “The SES was pretty good as long as the orientation was moved to reflect the robot's view and perspective.”
- “I loved the interface with the SES compared to the interface without it. For me this is because I play many video games and like to have more of a true 3D spatial knowledge of my surroundings, the interface without the SES had too

many blind spots and if I were trying to locate things it would be very difficult with so many blind spots. Yet I also picked up a moment when the SES failed to detect things due to shadows and once even it failed to pick things up that were out of range of the camera. If there was a way to make the camera move in a true 360 degree turn the SES would be so much more powerful and accurate with its data.”

### **Landmark Map**

The landmark map was employed to provide move to object and move to point commands. During the training, the participants saw the remote drive interface that is used to teleoperate the robot. During the experimental design, it was determined that the move to point and move to object commands would provide an easier interface. Several participants requested a manual drive mode because they felt it would have been helpful. Two participants stated that a joystick for driving would have been the most simple form of teleoperation. This addition would have also resolved requests by some participants to drive the robot in reverse. The participants also stated that it was difficult to make fine angle movements. One participant felt that the drive commands lacked flexibility because they did not have the remote drive.

A major source of frustration was the robot's odometry error. This error caused the robot position and map update to diverge greatly the more the robot moved. Another major source of frustration was the participants' confusion with their right and left versus the robot's right and left. Some participants resolved this dilemma by inverting their paper map to match the orientation of the map on the computer. The problem was that the robot was driving from north to south and the map was orientated such that north was up on the

computer screen. Many participants said they had to compensate for deficiencies in the interface and the error between the requested position and the robot's position.

Many participants stated that after some practice that the method for driving the robot was very easy. They also stated that the map with the landmarks greatly simplified the tasks. Some of the user comments included:

- “I was getting frustrated when the robot kept on going off course. I was also frustrated because I was having a hard time telling the colors of the cones apart. When I told the robot to travel to a certain point, if the point was closer to the robot it would get there more accurately than if the point was further away.”
- “It was frustrating running into the wall all the time. After figuring out what I was doing...it was not irritating at all. It just took practice which usually has to happen to be good at anything.”
- “The *move to point* command sequence seemed a little awkward. I would open the *move to* window, then have to click back on the map itself, and then go back and hit okay in the other window. It was also hard to reconcile the two. The one I looked at (paper one) was aligned one way, but on the screen it went the other way. Perhaps it'd be possible to invert the map on the computer screen?”

### **Robot Reactions**

During the evaluation, many participants stated that they were executing tasks with a mobile robot for the first time. In light of this, the participants' had very strong opinions about what they thought the robot should do. The most overwhelming comment was that the participants' wanted to drive the robot faster. The second most prevalent comment was

that they felt the robot needed an alignment because it frequently veered to the right and left while moving. One comment was,

“The robot needs an alignment. My frustration occurred when the robot did not drive straight. We had some difficulties that caused the test to run longer than anticipated but overall I enjoyed the test and learned many things about robots that I have not been exposed.”

### **Interface Reactions**

Some of the general comments regarding the interface related to the visual layout. Although the participants were informed that the interface was a prototype form, several suggested that it would be beneficial for all of the displays to be in one window as opposed to three separate windows. The camera, sonar, laser, and camera were in one window. The SES and landmark map had their own individual windows. One common statement was that the system had a slow reaction time and this was due to the wireless network communication with the robot.

In preparation for the human factors study, a prototype of the interface screens was developed. Each display component had several view options. These views were made into color printouts and a paper survey was performed with ten participants. The most prevalent display view selections were then used as the default view for the interface. Additionally, during the study orientation, each participant was asked to complete a pre-experiment questionnaire with choices for display preferences for the sonar, laser, compass, and SES. These participants completed this questionnaire without any training or explanation about what the views were. A couple of participants stated that their original choices for display

preferences would have really confused them after utilizing the dynamic interface. Figure 7.1 is the electronic survey for the display views.

One prevalent observation was that the interface was very visual, colorful, and easy to use. Participants felt that their confidence increased quickly as they time spent more time using the interface. Students also expressed displeasure with the amount of error inherent in some sensory data such as the sonar, laser, compass, map, and SES. A couple of participants suggested that a joystick or some sort of control pad to drive the robot would be very useful. Some examples of user comments were:

- “On the whole I thought the system was really good. Very easy to understand and use and I thought it had a very good response time.”
- “I think the interface would take a little practice or training before anyone would be an ace at it. I think after a couple tries anyone would rock at the tasks.”
- “The interface was easy but a little intimidating for the novice user.”

This chapter has presented the qualitative and quantitative results of the enhanced human-robot interface evaluation. These results included ten of the 27 participants, a discussion of system failures, statistics and user comments. The non-parametric statistics included the Kruskal-Wallis Rank Test, Wilcoxon Rank Sum Test and Friedman Test. There was also a non-parametric correlation statistic called the Spearman Rank Correlation. The quantitative data was user comments during the post-task and post-experiment questionnaire.

## CHAPTER VIII

### DISCUSSION / CONCLUSIONS / FUTURE WORK

The overall objective of this research was to determine if the addition of a Sensory EgoSphere (SES) enhanced a human-robot interface. The SES represents the short-term sensory memory of a mobile robot. It was hypothesized that the addition of the SES to a graphical based HRI would provide vital information to the participant concerning mission status and robot status, as well as the robot's present knowledge of itself. The purpose of this work was:

- To develop an enhanced agent-based human-robot interface with the addition of the SES agent,
- To conduct human factors experiments and evaluate the enhanced HRI, and
- To decrease the task execution time and situation assessment errors for potential users.

The research hypotheses were:

1. The SES decreases participant mental workload with the addition of a more intuitive display of sensory data.
2. The SES increases participant situational awareness of the robot status and the task/mission status.

In order to test the validity of these hypotheses, there were several research questions to be answered, that included: *Can the addition of the SES to a GUI,*

1. *Decrease task execution time?*
2. *Decrease the participants' workload?*

3. *Decrease errors in participant responses?*
4. *Improve the participants' situation awareness?*
5. *Improve the HRI's usability?*

## **Discussion**

This section discusses the results of the user study and their relationship to the study hypotheses, research questions, and the purpose.

### **Reduction in Task Execution Time**

In order to verify the first research question, task completion times were compared for the training tasks (A,B) and the teleoperation tasks (E, F). The fact that the second execution of a task with either the enhanced or the original interface generally produced a decrease in task completion time demonstrated that participants exhibited some level of learning. The overall training task time for the original interface was an average of 4 minutes (30.3%) shorter than for the enhanced interface, which included the SES. The only reduction in sub-task time for the enhanced interface was the driving directions. This is somewhat surprising, since this portion of the task did not involve necessarily using the interface. The driving directions time was approximately 10 seconds shorter for the enhanced interface.

The teleoperation tasks also demonstrated the same learning effect; the second run of a task was usually shorter. Additionally, a comparison of the original and enhanced interface showed that the original interface had an average decrease in the task time of 2 minutes (11.51%). The only significant result was the difference in the original interface completion time dependent on task presentation order. When original interface was

performed after the enhanced interface, it took a significantly less amount of time to complete.

In both sets of tasks, the total completion time was larger by at least 11% when the SES was added to the interface. Therefore, its addition actually increased completion times by 4 minutes and 2 minutes for the training and teleoperation tasks, respectively, rather than reduced them. Table 7.15 and Table 7.17 demonstrate the differences in the task completion times.

### **Reduction in Mental Workload**

In order to evaluate the first hypothesis, the MRQ and NASA-TLX methodologies were administered to each participant. The purpose of the MRQ evaluation was to determine if the enhanced interface reduced the amount of resources participants had to use to complete the task. The assumption was that reduced resources would imply a reduction in the perceived mental workload. It was shown in the previous chapter via the correlation analysis that a relationship between the resources and workload existed. There were correlations between some of the MRQ categories and the NASA-TLX, such as the frustration level, stress level, spatial attentive resource, spatial quantitative resource, spatial positional resource, necessary thinking, and mental demand.

In a comparison of participants' responses, a higher numerical value for a particular resource implied that the participant used that resource more to complete a task independent of task order. In a comparison of the enhanced and original interfaces, it was shown that the enhanced interface required fewer multiple resources. This was true for all categories except the spatial emergent. The spatial emergent category dealt with detecting an object from a highly cluttered or confusing background using vision. Since the enhanced interface included

the SES, this may have accounted for the increased usage of this resource for the enhanced interface. The overall MRQ result may also imply a reduction in mental workload for the enhanced interface.

The examination of the teleoperation tasks found a contradictory result. The manual resources were the same for the original and enhanced interfaces. The original interface had higher spatial quantitative and visual lexical resources. The remaining resource ratings were higher for the enhanced interface including the overall rating. The results of the MRQ disprove the concept of the enhanced interface reduction the multiple resources usage. The enhanced interface actually increased the demand on the multiple resources by approximately 5%. However, the training task did show a reduction in the demand on the resources by approximately 11%. One reason for this increase may be that the SES did not provide as much assistance when the robot was moving, therefore an increase in resource usage was found. The teleoperation tasks may have actually increased mental workload with the enhanced interface based upon the increased resources.

The hypothesis was that the addition of the SES to the original interface would reduce the participant's perceived mental workload. Higher perceived workload in this assessment indicates that participants assigned a higher numerical value, closer to 100, for a category if they used it more to complete a task. The enhanced interface for the training task demonstrated higher demands for the necessary thinking, task difficulty, physical effort, and stress level. These categories may have been higher because of the addition of the SES display. The additional display may have required more thinking, more use of the interface, more mouse clicks, and additional stress. The comparison of the original versus enhanced interface showed that there was a reduced overall perceived workload for the enhanced

interface. This result confirmed the research hypothesis for the enhanced interface. The difference in the performance satisfaction was statistically significant.

In a comparison of the original to the enhanced interface, it was found that the enhanced interface received higher ratings for the frustration and stress levels. It is believed that these are attributed to the odometry error as well as the SES display. The overall comparison of the enhanced and original interface showed that the perceived mental workload was reduced for the enhanced interface by approximately 13%. This result does indeed imply verification of the conjecture that the enhanced interface would reduce the perceived mental workload, but these results were not statistically significant.

In conclusion, the raw data implies confirmation of the hypothesis but the statistical analysis did not find a significant relationship. Since none of these results were significant, there is a need for additional testing. There was only one contradictory result and this was the MRQ analysis for the training tasks. This result was contradictory because although the workload was less, the participant used more resources for the enhanced interface.

### **Reduction in Participant Errors**

In order to verify the third research question, reduction in participant errors, task scores were compared for the training and teleoperation tasks. A higher score would imply that there were less participant errors. The original interface during the training task demonstrated higher task scores for the robot placement, robot orientation, cone placement, and cone color. The driving directions score was higher for the enhanced interface. Overall, the score for the original interface was higher than the score for the enhanced interface therefore the enhanced interface increased participant errors.

The teleoperation cone placement and overall scores were higher for the original than for enhanced interface. The cone color score, however, was higher for the enhanced interface. This comparison showed an improvement in the overall score of approximately two points when the enhanced interface was employed. Although this was not statistically significant, it appears that the enhanced interface may have slightly decreased participant errors.

In summary, the training tasks showed that the original interface had fewer participant errors while the enhanced interface had fewer errors during the teleoperation task. Therefore, the conjecture of the enhanced interface reducing participant errors was not shown.

### **Increased Situation Awareness**

In order to verify the second research hypothesis, the task sub-scores were evaluated for all tasks. The three levels of situation awareness are perception, comprehension and prediction. This work proposes that the addition of the SES to the HRI will move the participants' SA from the perception to comprehension level. Therefore, the SA was evaluated by examining certain task scores. For the two training tasks, the theory was that the cone color score may not be differentiated between the two tasks since this would be considered the perception level. However, it was thought that the robot placement, robot orientation, cone placement, and driving directions scores would improve for the enhanced interface. These scores correspond to the comprehension level. The results showed that the driving directions score was the only improved score with the enhanced interface. This improvement implies that the second hypothesis may be partially validated, although the results were not statistically significant.

With respect to the teleoperation tasks, the cone color score was again the perception level of SA. The cone placement score should have been improved with the addition of the SES to the interface. The results showed that there was an average of a 21-point improvement for the cone placement score with the enhanced interface. Therefore, it was implied that the hypothesis of improved situation awareness for the enhanced interface was shown for the cone placement task although it was not statistically significant.

In summary, the hypothesis of increased situation awareness could only be partially proven for the training task. The raw data implies confirmation of the SA hypothesis for the training and teleoperation tasks although it was not statistically significant. This result demonstrates a need for further testing for absolute confirmation.

### **Improved Usability**

In order to evaluate the fifth research question concerning improved HRI usability, several questionnaires were administered to the participants. These questionnaires included four post-task and one post-experiment questionnaire.

In a comparison of the original versus enhanced interfaces for the training task, the enhanced interface task was rated as easier, clearer, more stimulating, and satisfying. Therefore, the enhanced interface was shown to be slightly more usable for the task specific ratings.

With respect to the general ratings, the enhanced interface task was rated higher for the ability to understand and interpret data, correct errors, control the system, and complete tasks. Once again, the results for the general task ratings were that the enhanced interface provided better usability.

The system capability ratings were evenly divided between the original and enhanced interfaces. Participants rated the original interface higher for power level, ease of task completion, and support information satisfaction. The enhanced interface was rated higher for the system speed, flexibility, and task completion time. Therefore, the assumption of improved usability for the enhanced interface could not be confirmed for the system capability ratings.

The comparison of the enhanced and original interfaces task specific ratings showed that the enhanced interface task was rated as easier, clearer, and more stimulating. The original interface for the teleoperation task was rated as more satisfying. These results indicate that the enhanced interface task was easier, clearer and more stimulating although it was also slightly more frustrating. Therefore, the enhanced interface was shown more usable over most of the task specific ratings.

The original interface for the teleoperation task was rated higher for the ability to understand and interpret data, correct task errors, and control the system. These results indicate that the enhanced interface was shown to be more usable for the general ratings.

The results of the system capability ratings were similar to those found for the training tasks. The original interface for the teleoperation task was rated higher for the system speed, flexibility, ease of task completion, and support information satisfaction. The enhanced interface task was rated higher for the power level and task completion time. Therefore, the results were split between the enhanced and original interface. The assumption of improved usability could not be proven for the system capability ratings.

In summary, the enhanced interface was shown to increase the usability for task specific and general ratings. It was not shown to increase the general system usability. The reason for the latter result could be attributed to the distributed windows environment, the

wireless network, or the time delay when attempting to display such a high amount of graphic information.

The post-experiment questionnaire was employed to obtain overall usability results for the HRI. In general, the system was rated as wonderful, easy to use, stimulating and flexible. Participants were satisfied with the overall system but the system power was perceived as requiring improvement as indicated by the first set of post-experiment questions.

The question set related to the display screen information and highlighted several areas for improvement. The organization of information was clear, the ability to read characters was easy, and the buttons were easy to use. The display also had a simple and natural dialogue, and it was easy to locate necessary task information. Participants also felt the items were grouped for direct comparison, the active window was clear, background contrast was excellent, the cursor and menu were distinct, and the screen sequencing was clear. The definition of exit markings is that they are indications of how to close the interface. The definition of controls and action compatibility is that certain user inputs or display devices correspond closely to the action performed on the robot. The areas that required improvement are:

- the addition of clearly marked exits,
- increased controls and action compatibility,
- display layout modification in order to simplify tasks, and
- reduce the display clutter.

The question set that probed system features showed that the use of terms was consistent, message positions were consistent, input prompts were clear, and the computer

always informs the user of its progress. Some areas that require improvement are the addition of help and documentation, as well as error messages.

The question set regarding the ability to learn the system showed that learning to operate the system was easy. The time to learn the system was fast and remembering how to enter commands was easy. Additionally, participants felt tasks could be performed in a straightforward manner. Necessary improvements include:

- Improve system functionality and capability,
- Decrease the number of steps to perform tasks, and
- Creation of keyboard shortcuts.

Results indicated that all system capability categories require improvement except for the ability to easily correct mistakes as well as the ability of novices to accomplish tasks with proper training. The improvements include:

- Increase system speed,
- Increase response time,
- Increase information display rate, and
- Decrease system failures.

All ratings were found to be acceptable for the overall system. Therefore, no needed improvements were found. Participants agreed that they liked using the system. Participants agreed that the system interface was pleasant. Participants agreed that they felt satisfied with the overall system. Participants strongly agreed that it was easy to learn the system and that they felt comfortable using the system.

In conclusion, the mental workload hypothesis was partially proven for the enhanced interface for the training and teleoperation tasks. This result was found based upon the reduction in some of the MRQ multiple resource ratings for the enhanced interface. The

increased situation awareness was partially confirmed for the training tasks since improvement from the perception to the comprehension level was shown for some sub-scores. The increased situation awareness was partially proven for the teleoperation task since the cone placement score did improve for the enhanced interface. Since the results were not statistically significant, neither hypothesis could be proven. Further testing would be required to fully substantiate the hypotheses.

### **Conclusions**

In conclusion, although the user study did not fully support the research hypotheses or confirm several research questions, it did produce some very intriguing results. The large number of significant correlations between several of the variables demonstrates a need for a second phase of evaluations. The second phase of evaluations will be performed using the original HRI and a revised enhanced HRI. The results provide a low-level confirmation of the hypotheses and research questions. The need for additional evaluations is based upon the fact that although some results showed change, they were not statistically significant. The reduction in task time and reduction in participant errors could not be confirmed. The reduction in perceived mental workload and increased situational awareness were partially confirmed. These results along with the usability evaluation suggest modifications to the interface and SES display. The evaluations should be rerun with tasks that are more stringent on a larger sample size. Additionally, some of the other influences, as indicated by the correlation analysis, on the workload, task time, and task score should be minimized. This thesis has presented the implementation of an enhanced HRI with the addition of the SES. The interface was implemented using Visual Basic, OpenGL, and IMA. The interface was evaluated using 27 participants performing four tasks but the data analysis was

performed with only ten participants. The data collected included: the number of mouse clicks, task completion time, task score, usability, and perceived mental workload. A statistical analysis was performed using the data from the ten participants who completed both of the teleoperation tasks with no major system or hardware failures. The non-parametric analyses included the Spearman rank correlation and the Kruskal-Wallis rank test. These results were analyzed in order to attempt to determine the validity of the research hypotheses.

### **Future Work**

In addition to the aforementioned second phase user evaluation, the qualitative results highlighted some SES display improvements. One possible improvement relates to the SES shape. This change would transform the perfect sphere to an oval, thus stretching the SES into more of a football type shape. Such a shape may minimize the necessity to manipulate the SES as frequently in order to view the node postings. A second improvement would add compass and odometry information to the SES display in order to support the correlation of the SES display with the robot's movement and heading. The odometry link should enable objects to stream on the SES display as the robot moves. For example, as the robot moves forward, posted objects move towards the back of the sphere and eventually disappear after a certain time decay factor.

In a higher level functionality, the participant would also have a link to the short-term memory database and use the SES to provide graphical displays of the robot's memory at certain time instances. The purpose of this function would be to provide greater utility with an SES that can be used to view past events. The SES would provide information from

some defined time interval from the past. The display would then update the view based upon the data time stamp stored in the database.

In addition, because the SES is so graphically intensive, to aid in real-time updates, it may be necessary to change from distributed Windows-based machines to a main frame system or higher-powered personal computers. The justification would be that the system would run faster therefore improving the system speed and response time ratings. Additionally, this would aid in the SES data streaming as it moves forward in space since this requires continuous calculation of node postings for objects.

Finally, some HRI-based changes were indicated by the participants' comments. One change would be to create more dynamic sonar and laser displays linked to the robot's base. Such displays would move with the robot. The interface would also provide either a joystick or more flexible manual drive user interface during teleoperation. This along with the move to point and move to object behaviors would provide the capability of moving the robot in more directions.

Participants' comments indicated that there was an overwhelming amount of frustration with the robot's odometry error. One possible solution would be to periodically correct the robot's position based upon known landmarks or GPS, if outdoors.

In order to element some confusion with the SES graphic, the amount of training time could be doubled. This may also eliminate the apprehension of participants to use the display. The addition of the robot's backward facing camera view to the interface would resolve many participant issues with the limited view. Additionally, since the separation of the sonar and laser confused some participants, they could be consolidated onto one display as they are shown on the SES. Also, the sonar and laser data also could be added to the camera display. Participants felt the ability to determine distances on the SES was difficult

so the sonar, laser, and compass data could be added to the SES display to potentially relieve this difficulty. The addition a color correction algorithm would resolve issues of objects being posted with the wrong color. This usually happened due to lighting conditions, therefore a color correction algorithm would minimize the problem. The SES was also thought to be more helpful if it moved to correspond to the robot's orientation. This implies a linkage between the robot's base or compass that dynamically moves with the robot. Although, this capability could be added, the participant could also have the ability to move the display independent of the robot sensors.

With respect to the post task questions, some areas that demonstrated a need for improvement were the difficulty and frustration levels. A solution for these would be to increase training and modify the system to be more intuitive and user-friendly. As previously stated, the addition of faster computers or running the system on a mainframe may improve some of the system speed and power level issues. The addition of more innovative methods to interact with the robot through the interface may also improve the participants' rating of the time required to complete tasks. More detailed task descriptions would resolve issues with clarity level. Some of the aforementioned improvements would decrease the difficulty and frustration levels. The system's power capability would be improved with the addition of better computers. Since the interface is a prototype, it did not contain any exit markings. By consolidating some display screens and rearranging the remaining windows, it may be possible to improve the display layout to simplify tasks and reduce display clutter. Since this was a prototype, there was no help documentation. An improvement would be the creation of a help menu and interface documentation. There were no keyboard shortcuts, the addition of shortcuts might improve usability for intermediate to expert users.

## **Appendices**

### **A. NASA-TLX**

The NASA-Task Load Index (TLX) was given to each participant as part of the post-task questionnaire [Hart et al., 1998]. It was used to measure the perceived mental workload of participants during the execution of tasks. There was a combination of questions from the NASA-TLX, RNASA-TLX and the MRQ questionnaires [Cha et al., 2001]. This Appendix presents the NASA-TLX tool.

The NASA-TLX is a subjective workload rating each participant completed. The rating measures the amount of mental effort required by a participant to perform a given task. The task load index consists of six scales that are weighted to reflect the contribution of each factor to the overall workload. The weights are set by the participant and therefore represent their perspective of the workload effort. Table A.1 list these six components. This Appendix also contains the questionnaire given to each participant. For the purpose of this work, the overall rating for the workload was determined by averaging the sub-scales.

The NASA-TLX was given electronically to each participant and each scale was a horizontal slider bar with a title and opposite rating descriptions on each end. There was not any numerical information display and the participant rated each rating by sliding the bar. The bar represented numbers from 1 to 100 and each participants component value was weighted by the participants' weighting of the importance of that particular measure. Table A.1 is the task load subscales of the NASA-TLX [Hart et al., 1988; Cha et al., 2001]. The participants employed an electronic version of the NASA-TLX questionnaire shown in Appendix A.1. Appendix A.2 is a graphical presentation of the participants' results.

Table A.1. Task Load subscales of the NASA-TLX [Cha et al., 2001]

| Task Load Component | Description                                                                                                                            | Rating       | Code |
|---------------------|----------------------------------------------------------------------------------------------------------------------------------------|--------------|------|
| Mental Demand       | What amount of thinking was necessary to complete the task?                                                                            | Low.....High | MD   |
| Physical Demand     | Was the amount of physical activity was exhausting or not too strenuous?                                                               | Low.....High | PD   |
| Temporal Demand     | What amount of pressure did you feel due to the rate at which the task elements occurred?                                              | Low.....High | TD   |
| Performance         | How successful were you in doing what you were asked to do and how satisfied you were with what was accomplished?                      | Low.....High | OP   |
| Frustration         | How insecure, discouraged irritated, annoyed versus secure, gratified, content, and complacent did you feel while performing the task? | Low.....High | FR   |
| Effort              | What amount of mental and/or perceptual activity and physical effort was required to complete the task?                                | Low.....High | EF   |

There also exists an RNASA-TLX (Revision of NASA-TLX) that was created as a solution to problems discovered with the NASA-TLX [Cha et al., 2001]. Some of the problems include understanding and rating the six subscales of the NASA-TLX. The descriptions of the scales were vague and technical and it was difficult for participants to relate them to the particular task performed. For example, the RNASA-TLX for an in-vehicle navigation system would contain specific references to the task performed. Table A.2 is an example of the six subscales for the RNASA-TLX [Cha et al., 2001].

Table A.2. Task Load subscales of the RNASA-TLX

| Task Load Component                     | Description                                                                                                                                                                      | Rating       |
|-----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| Mental Demand                           | How much mental attention was needed during driving when using the navigation system?                                                                                            | Low.....High |
| Visual Demand                           | How much visual activity was required during driving when using a navigation system to recognize the information from a navigation system or other external information sources? | Low.....High |
| Auditory Demand                         | How much auditory activity was required during driving when using a navigation system to recognize or hear information presented from a navigation system?                       | Low.....High |
| Temporal Demand                         | How much time pressure was required due to rate or pace at the task elements occurred during driving using a navigation system?                                                  | Low.....High |
| Difficulty in Driving                   | How hard was it to drive when using a navigation system with other in-vehicle control equipment.                                                                                 | Low.....High |
| Difficulty in Understanding Information | How hard was it to understand information presented from the navigation system?                                                                                                  | Low.....High |

The rating categories from Tables A.1 and A.2 were both used on an electronic questionnaire to measure workload ratings for the training and evaluation tasks. The participants completed the electronic version of the NASA-TLX shown in Appendix A.1.

## A.1

### NASA-TLX

#### INSTRUCTIONS:

Please place a mark on the scale that represents the level to which you feel the given demand value was used to complete the task you just completed.

#### EXAMPLE:

The amount of stress experienced while completing this task was

LOW \_\_\_\_\_x\_\_\_\_\_ HIGH

#### Mental Demand

1. The amount of thinking necessary to complete this task was

LOW \_\_\_\_\_ HIGH

2. The task difficulty was

LOW \_\_\_\_\_ HIGH

#### Physical Demand

3. The amount of physical activity necessary to complete the task was

LOW \_\_\_\_\_ HIGH

#### Temporal Demand

4. The amount of time required to complete the task was

LOW \_\_\_\_\_ HIGH

5. The amount of time pressure I felt to complete the task was

LOW \_\_\_\_\_ HIGH

#### Performance

6. The level to which the goals set by the experimenter were achieved was

LOW \_\_\_\_\_ HIGH

7. The satisfaction felt with the performance during the execution of the task was

LOW \_\_\_\_\_ HIGH

**Effort**

8. The mental effort necessary to complete the task was

LOW \_\_\_\_\_ HIGH

9. The physical effort necessary to complete the task was

LOW \_\_\_\_\_ HIGH

**Frustration**

10. The level of frustration felt while executing the task was,

LOW \_\_\_\_\_ HIGH

11. The amount of stress and irritation felt while completing the task was

LOW \_\_\_\_\_ HIGH

## **A.2**

### **Graphical Presentation of Participants' Results**

## **B. MULTIPLE RATINGS QUESTIONNAIRE (MRQ)**

The MRQ is a 17-item subjective workload assessment used to measure the perceived workload of a participant based upon the multiple ratings theory [Boles et al., 2001a, 2001b]. The development of this questionnaire is described in Chapter II. The primary purpose of the MRQ is to identify the subjective mental workload on specific mental ratings. Unlike the NASA-TLX, it does not require the participant to use a weighting system. Some questions from this rating scale exam were administered to the participant in the post-task and post-experiment questionnaires. The participants completed the electronic version of the MRQ provided in Appendix B.1. Appendix B.2 is a graphical presentation of the participants' responses.

## B.1

### Multiple Ratings Questionnaire

The purpose of this questionnaire is to characterize the nature of the mental processing used in the task you just completed. Below are the names and descriptions of several mental processes. Please read each carefully so that you understand the nature of the processes. After reading the question, rate the task on the extent to which it uses each process. Please use the following scale to rate each of these tasks

| No usage | light usage | moderate usage | heave usage | Extreme usage |
|----------|-------------|----------------|-------------|---------------|
| 0        | 1           | 2              | 3           | 4             |

1. Manual process - Movement of arms, hands and/or fingers.
2. Short-term memory process - Remember information for a period of time ranging from a couple of seconds to half a minute.
3. Spatial attentive process - Focus of attention on a location using the sense of vision.
4. Spatial categorical process - Judgment of simple left-versus-right or up-versus-down relationships, without consideration of precise location using the sense of vision.
5. Spatial emergent process - Picking out a form or object from a highly cluttered or confusing background using the sense of vision.
6. Spatial positional process - Recognition of a precise location as differing from other locations using the sense of vision.
7. Spatial quantitative process - Judgment of numerical quantity based on a nonverbal, non-digital representation using the sense of vision.
8. Visual lexical process - Recognition of words, letter, or digits using the sense of vision

9. Visual temporal process - Judgment of time intervals, or the timing of events using the sense of vision

### **C. VANDENBERG MENTAL ROTATION TEST**

The Vandenberg mental rotation test is used to assess the spatial reasoning skills of an individual [Vandenberg et al., 1979; Shepard et al., 1971]. The following test was given to all participants during the orientation. The test contains six pages and two parts. The first two pages explain the format of the test, provides instructions for marking the answers and presents a sample problem. The next two pages are part of the exam. Ten questions were completed in three minutes. The final two pages were the second set of ten questions that the participant completed in three minutes. Each question is associated with a figure and four images. Two of the images are the figure rotated. The other two images may be a mirror image of the figure or a completely different figure. Therefore, there are two correct answers for each question. A question was scored as correct if both of the rotations of the figure were marked. If only one correct image was marked then the question was scored as incorrect. The score was calculated as the total number of correct responses. The participants completed the electronic version of the Vandenberg mental rotation test given in Appendix C.1. Appendix C.2 is the graphical presentation of participants' results.

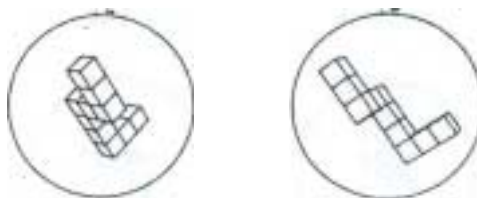
## C.1

### Vandenberg Mental Rotation Test

This is a test of your ability to look at a drawing of a given object and find the same object within a set of dissimilar objects. The only difference between the original object and the chosen object will be that they are presented at different angles. An illustration of this principle is given below, where the same single object is given in five different positions. Look at each of them to satisfy yourself that they are only presented at different angles from one another.

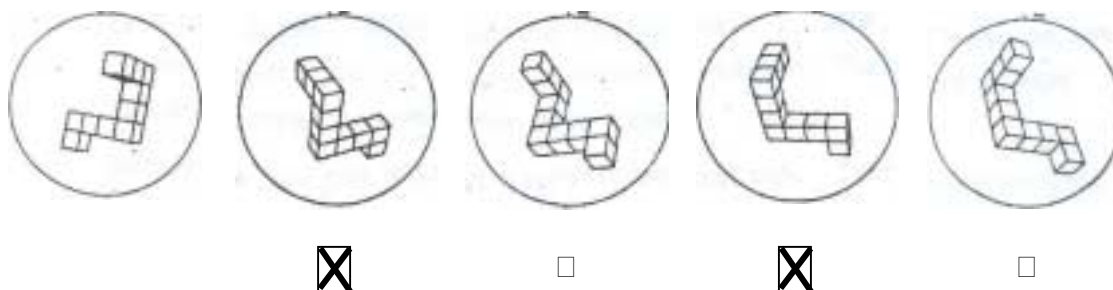


Below are two drawings of new objects. They cannot be made to match the above five drawings. Please note that you may not turn over the objects. Satisfy yourself that they are different from the above.

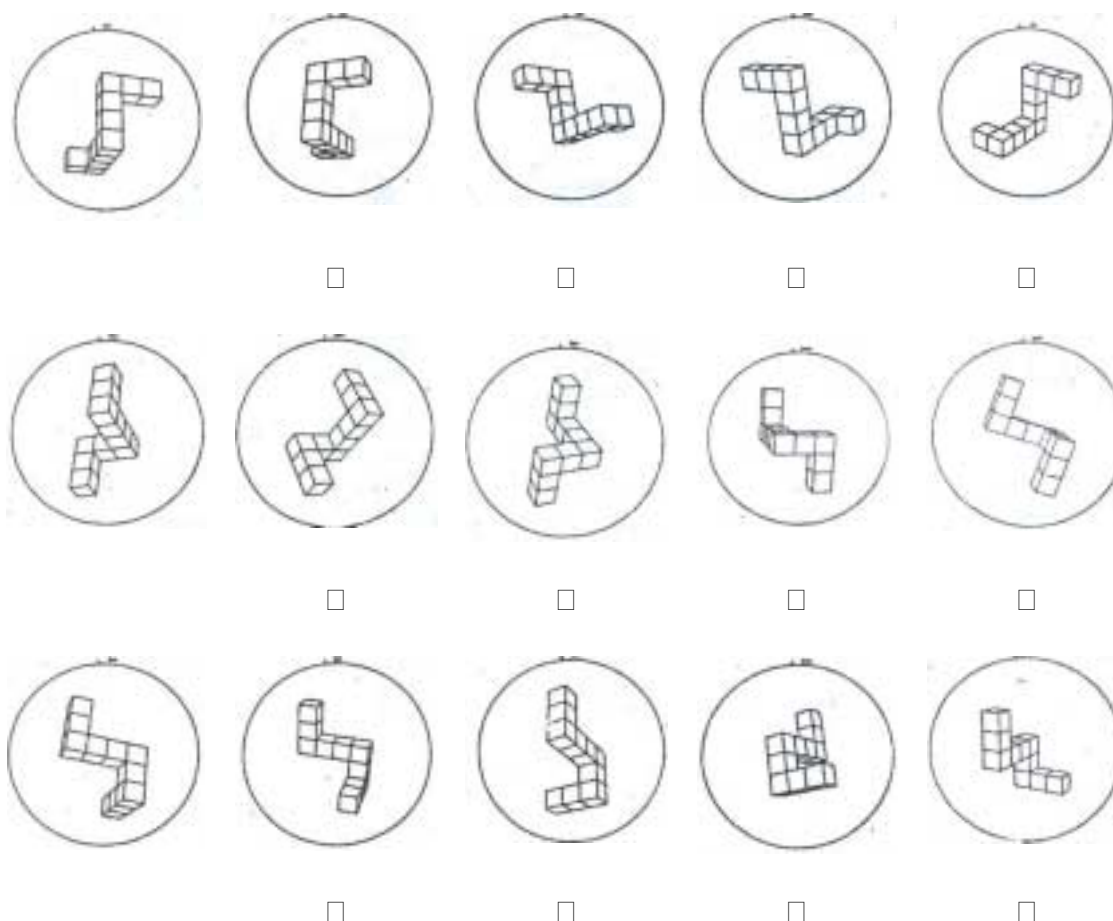


Now let's do some sample problems. For each problem there is a primary object on the far left. You are to determine which two of four objects to the right are the same object given on the far left. In each problem, always two of the four drawings are the same object as the

one on the left. You are to put 'X' in the boxes below the correct ones, and leave the incorrect ones blank. The first sample is done for you.



Do the rest of the sample problems yourself. Which two drawings of the four on the right show the same object as the one on the left? There are always two and only two correct answers for each problem. Put an X under the two correct drawings.



- ANSWERS: (1) first and second drawings are correct  
(2) first and third drawings are correct  
(3) second and third drawings are correct

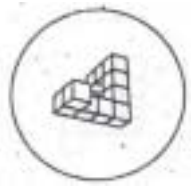
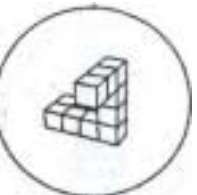
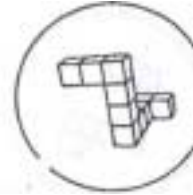
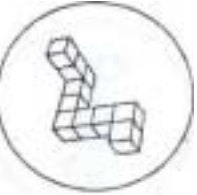
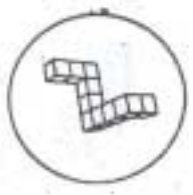
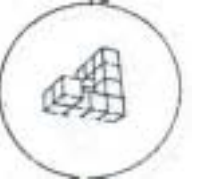
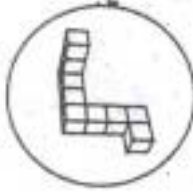
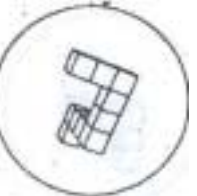
This test has two parts. You will have 3 minutes for each of the two parts. Each part has two pages. When you have finished Part I, STOP. Please do not go on to Part 2 until you are asked to do so. Remember: There are always two and only two correct answers for each time.

Work as quickly as you can without sacrificing accuracy. Your score on this test will reflect both the correct and incorrect responses. Therefore, it will not be to your advantage to guess unless you have some idea which is correct.

---

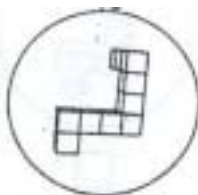
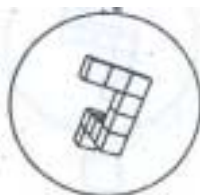
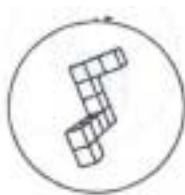
STOP. Do not turn this page until asked to do so

# PART I

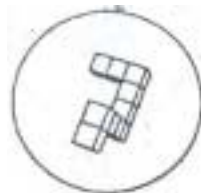
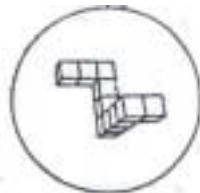
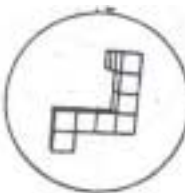
1.       
☐ ☐ ☐ ☐
2.       
☐ ☐ ☐ ☐
3.       
☐ ☐ ☐ ☐
4.       
☐ ☐ ☐ ☐
5.       
☐ ☐ ☐ ☐

Go to the next page

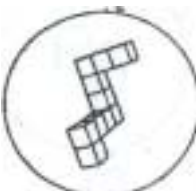
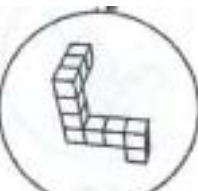
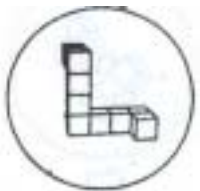
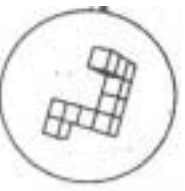
6.


☐
☐
☐
☐

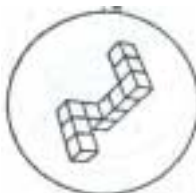
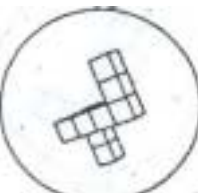
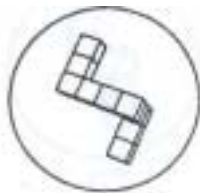
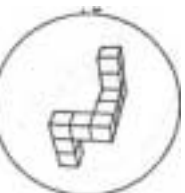
7.


☐
☐
☐
☐

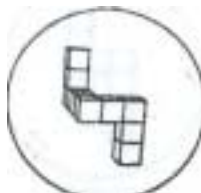
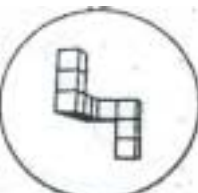
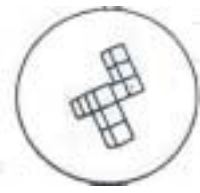
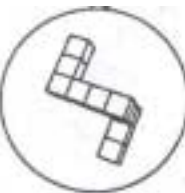
8.


☐
☐
☐
☐

9.


☐
☐
☐
☐

10.


☐
☐
☐
☐

STOP. Do not turn this page until asked to do so

## PART II

11.


☐
☐
☐
☐

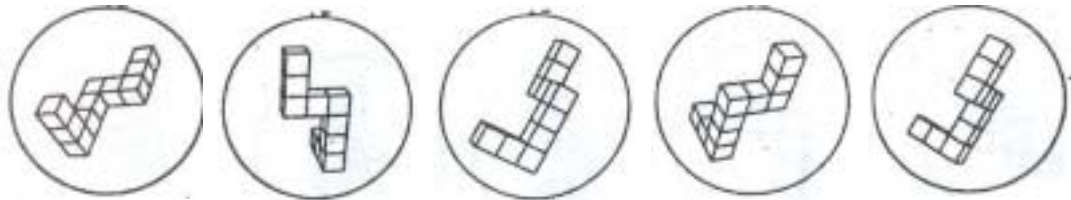
12.


☐
☐
☐
☐

13.


☐
☐
☐
☐

14.

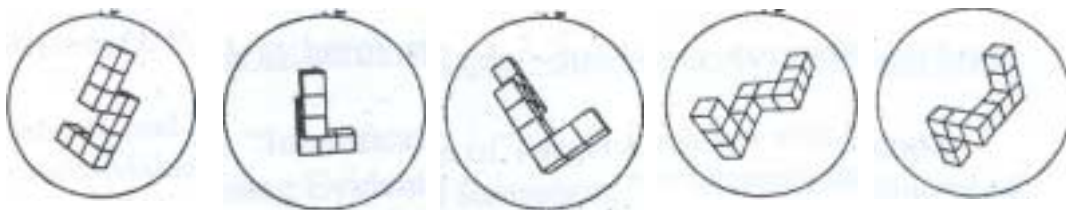

☐
☐
☐
☐

15.

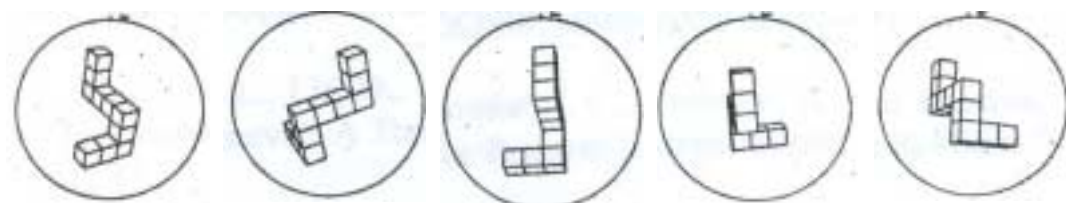

☐
☐
☐
☐

Go to the next page

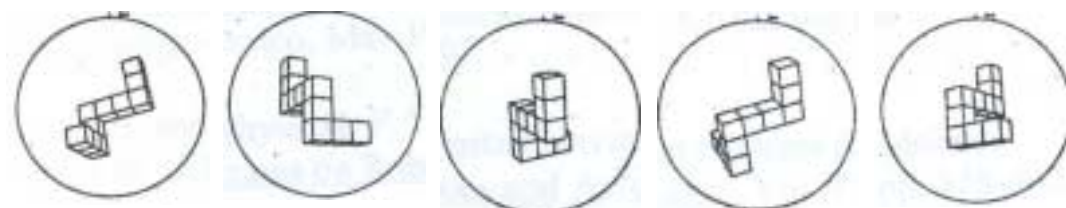
16.



17.



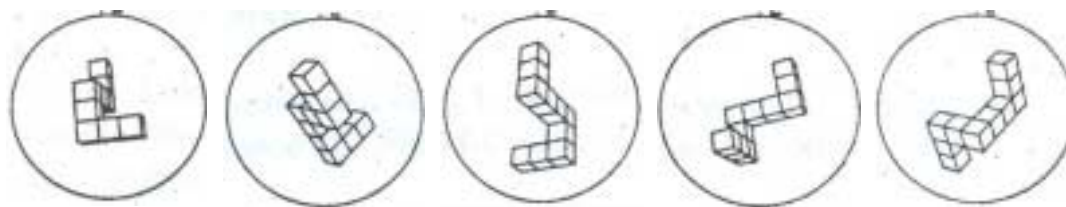
18.



19.



20.



STOP.

## **C.2**

### **Graphical Presentation of Participants' Results**

## **D. PRE-EXPERIMENT QUESTIONNAIRE**

The follow questionnaire was administered to each participant during the orientation before the first task was assigned. The purpose of the questionnaire was to understand each participants' demographic background as well as ensure that each participant met the participation requirements. This questionnaire and the spatial reasoning test were administered on the same day [Adams, 1995; Dumas et al., 1999; Rosson et al., 2002]. Each participant completed the electronic version of the pre-experiment questionnaire provided in Appendix D.1. Questions 7 through 11 were rated using a horizontal slider bar with values from 1 to 100 representing the low to high scale. Appendix D.2 is the graphical presentation of participants' responses.

## D.1

### Pre-Experiment Questionnaire

\*\*\*\*\*PERSONAL\*\*\*\*\*

Name: \_\_\_\_\_

Gender: Male/Female

Age: \_\_\_\_\_

Occupation: \_\_\_\_\_

Major (if student): \_\_\_\_\_

Highest level of education: Some HS/HS/ Associates/Bachelors/Masters/Doctorate

\*\*\*\*\*COMPUTER EXPERIENCE\*\*\*\*\*

1. For how many years have you been using personal computers?  
\_\_\_\_\_ (years)
2. Do you use personal computers for task such as word processing or spreadsheets?  
\_\_\_\_\_ (yes/no)
3. How many times a week do you use a computer? \_\_\_\_\_ (days)
4. How many times a day do you use a computer? \_\_\_\_\_ (hours)
5. How many times a week do you play video games? \_\_\_\_\_ (days)
6. How many times a day to you play video games? \_\_\_\_\_ (hours)

For the following questions, please mark an “X” on the line at the point that you feel adequately reflects your experience.

7. What level of experience do you have using computers?

LOW \_\_\_\_\_ HIGH

8. What is your level of experience with computer graphics?

LOW \_\_\_\_\_ HIGH

9. What is your level of experience with video games?

LOW \_\_\_\_\_ HIGH

10. What is your level of experience with robots?

LOW \_\_\_\_\_ HIGH

11. What is your level of experience with mobile robots?

LOW \_\_\_\_\_ HIGH

12. What is your level of experience with teleoperation of mobile robots?

LOW \_\_\_\_\_ HIGH

## **E. POST-TASK QUESTIONNAIRE**

The post-task questionnaire was administered to each participant upon the completion of the given task. The purpose of the questionnaire was to measure mental workload, user satisfaction with the interface and usability. These questions were based upon the SWAT, MCH, QUIS and MRQ questionnaires. Along with this questionnaire, the participant also completed NASA-TLX measurement described in Appendix A. A detailed explanation of these questionnaires can be found in Chapter II [Adams, 1995, Chin et al., 1988, Davis, 1989, Dumas et al., 1999, Hix et al., 1993, Lewis, 1995, Lin et al., 1997, Nash, 2000, Nielsen, 1993, Ravden et al., 1989, Shneiderman, 1987]. The participants completed the electronic version of the post-task questionnaire given in Appendix E.1. Appendix E.2 is the graphical presentation of participants' responses.

## E.1

### Post-Task Questionnaire

Please rate your experiences with this system. Try to respond to all of the questions. If the item is not applicable, please mark N/A. Place a “X” under a value from 0 to 9 or under the N/A column. There will be room at the end of the questionnaire to add comments.

1. This task was

|    | 1                | 2                    | 3       | 4                    | 5                | 0   |
|----|------------------|----------------------|---------|----------------------|------------------|-----|
| a. | Very difficult   | Slightly difficult   | Neutral | Slightly easy        | Very easy        | N/A |
| b. | Very Confusing   | Slightly Confusing   | Neutral | Slightly Clear       | Very Clear       | N/A |
| c. | Very Dull        | Slightly Dull        | Neutral | Slightly Stimulating | Very Stimulating | N/A |
| d. | Very Frustrating | Slightly Frustrating | Neutral | Slightly Satisfying  | Very Satisfying  | N/A |

2. I was able to interpret and understand data readings during the task

| 1     | 2            | 3       | 4          | 5      | 0   |
|-------|--------------|---------|------------|--------|-----|
| Never | Infrequently | Neutral | Frequently | Always | N/A |

3. I was able to correct my errors during the task.

| 1      | 2         | 3       | 4                | 5      | 0   |
|--------|-----------|---------|------------------|--------|-----|
| Rarely | Sometimes | Neutral | Most of the Time | Always | N/A |

4. I felt in control of the system during the task.

| 1      | 2         | 3       | 4                | 5      | 0   |
|--------|-----------|---------|------------------|--------|-----|
| Rarely | Sometimes | Neutral | Most of the Time | Always | N/A |

5. I felt able to complete the task.

| 1      | 2         | 3       | 4                | 5      | 0   |
|--------|-----------|---------|------------------|--------|-----|
| Rarely | Sometimes | Neutral | Most of the Time | Always | N/A |

6. The system's capabilities for this task were

|    | 1                | 2                         | 3       | 4                     | 5              | 0   |
|----|------------------|---------------------------|---------|-----------------------|----------------|-----|
| a. | Too Slow         | Slightly Slow             | Neutral | Almost Fast Enough    | Fast Enough    | N/A |
| b. | Inadequate Power | Slightly Inadequate Power | Neutral | Almost Adequate Power | Adequate Power | N/A |
| c. | Rigid            | Slightly Rigid            | Neutral | Slightly Flexible     | Very Flexible  | N/A |

7. Overall, I am satisfied with the ease of completing tasks with this interface.

| 1                 | 2        | 3       | 4     | 5              | 0   |
|-------------------|----------|---------|-------|----------------|-----|
| Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree | N/A |

8. I am satisfied with the amount of time it took me to complete the task.

| 1                 | 2        | 3       | 4     | 5              | 0   |
|-------------------|----------|---------|-------|----------------|-----|
| Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree | N/A |

9. Overall, I am satisfied with the support information when completing the task.

| 1                    | 2        | 3       | 4     | 5                 | 0     |
|----------------------|----------|---------|-------|-------------------|-------|
| Strongly<br>Disagree | Disagree | Neutral | Agree | Strongly<br>Agree | N / A |

Please write your comments to the on the back of this page.

## **F. POST-EXPERIMENT QUESTIONNAIRE**

The follow questionnaire was administered to each participant upon completion of all assigned tasks, during the last evaluation session. The purpose of this questionnaire was to obtain the participants' overall rating of the two human-robot interfaces and any final comments [Adams, 1995, Chin et al., 1988, Davis, 1989, Dumas et al., 1999, Hix et al., 1993, Lewis, 1995, Lin et al., 1997, Nash, 2000, Nielsen, 1993, Ravden et al., 1989, Shneiderman, 1987]. Participants completed the electronic version of the post-experiment questionnaire provided in Appendix F.1. Appendix F.2 is the graphical presentation of participants' responses.

## F.1

### Post-Experiment Questionnaire

Please rate your satisfaction with the system. Try to respond to all of the questions.

If the item is not applicable, please mark NA. Place a “X” under a value from 0 to 9 or under the N/A column. There will be room at the end of the questionnaire to add comments.

#### 1. Overall Reaction to the Software

|    | 1                     | 2                         | 3       | 4                       | 5                | 0   |
|----|-----------------------|---------------------------|---------|-------------------------|------------------|-----|
| a. | Terrible              | Slightly Terrible         | Neutral | Slightly Wonderful      | Wonderful        | N/A |
| b. | Very Difficult        | Slightly Difficult        | Neutral | Slightly Easy           | Very Easy        | N/A |
| c. | Very Frustrating      | Slightly Frustrating      | Neutral | Slightly Satisfying     | Very Satisfying  | N/A |
| d. | Very Inadequate Power | Slightly Inadequate Power | Neutral | Slightly Adequate Power | Adequate Power   |     |
| e. | Very Dull             | Slightly Dull             | Neutral | Slightly Stimulating    | Very Stimulating | N/A |
| f. | Very Rigid            | Slightly Rigid            | Neutral | Slightly Flexible       | Very Flexible    | N/A |

#### 2. Screen

|                                     | 1         | 2             | 3       | 4             | 5         | 0   |
|-------------------------------------|-----------|---------------|---------|---------------|-----------|-----|
| a. reading characters on the screen | Very Hard | Slightly Hard | Neutral | Slightly Easy | Very Easy | N/A |

|                                                               |                |                    |         |                      |                  |       |
|---------------------------------------------------------------|----------------|--------------------|---------|----------------------|------------------|-------|
| b. organization of the information                            | Very Confusing | Slightly Confusing | Neutral | Slightly Clear       | Very Clear       | N / A |
| c. screen window layout                                       | Very Unhelpful | Slightly Unhelpful | Neutral | Slightly Helpful     | Very helpful     | N / A |
| d. use of command buttons                                     | Very Difficult | Slightly Difficult | Neutral | Slightly Easy        | Very Easy        | N / A |
| e. there were clearly marked exits                            | Never          | Almost Never       | Neutral | Almost Always        | Always           | N / A |
| f. interface had simple and natural dialogue                  | Never          | Almost Never       | Neutral | Almost Always        | Always           | N / A |
| g. there is a good compatibility between controls and actions | Never          | Almost Never       | Neutral | Almost Always        | Always           | N / A |
| h. finding the information necessary to complete the task     | Very Difficult | Slightly Difficult | Neutral | Slightly Easy        | Very Easy        | N / A |
| i. display layout simplifies task                             | Never          | Almost Never       | Neutral | Almost Always        | Always           | N / A |
| j. display                                                    | Very Cluttered | Slightly Cluttered | Neutral | Slightly Uncluttered | Very Uncluttered | N / A |

|                                                |                |                    |         |                    |                |       |
|------------------------------------------------|----------------|--------------------|---------|--------------------|----------------|-------|
| k. items are paired for direct comparison      | Very Confusing | Slightly Confusing | Neutral | Slightly Clear     | Very Clear     | N / A |
| l. the active window is indicated              | Very Confusing | Slightly Confusing | Neutral | Slightly Clear     | Very Clear     | N / A |
| m. character contrast with background          | Very Poor      | Slightly Poor      | Neutral | Slightly Excellent | Very Excellent | N / A |
| n. the cursor is distinctive                   | Very Confusing | Slightly Confusing | Neutral | Slightly Clear     | Very Clear     | N / A |
| o. the display elements are distinctive        | Very Confusing | Slightly Confusing | Neutral | Slightly Clear     | Very Clear     | N / A |
| p. groups of items demarcated                  | Very Confusing | Slightly Confusing | Neutral | Slightly Clear     | Very Clear     | N / A |
| q. menus are distinctive from other data items | Very Confusing | Slightly Confusing | Neutral | Slightly Clear     | Very Clear     | N / A |
| r. sequence of screens                         | Very Confusing | Slightly Confusing | Neutral | Slightly Clear     | Very Clear     | N / A |

### 3. Terminology and System Information

|                                       | 1                 | 2                     | 3       | 4                   | 5               | 0     |
|---------------------------------------|-------------------|-----------------------|---------|---------------------|-----------------|-------|
| a. use of terms throughout the system | Very Inconsistent | Slightly Inconsistent | Neutral | Slightly Consistent | Very Consistent | N / A |

|                                        |                   |                       |           |                     |                 |       |
|----------------------------------------|-------------------|-----------------------|-----------|---------------------|-----------------|-------|
| b. position of messages on screen      | Very Inconsistent | Slightly Inconsistent | Neutral   | Slightly Consistent | Very Consistent | N / A |
| c. prompts for input                   | Very Confusing    | Slightly Confusing    | Neutral   | Slightly Clear      | Very Clear      | N / A |
| d. computer informs about its progress | Never             | Almost Never          | Sometimes | Almost Always       | Always          | N / A |
| e. help and documentation              | Very Unhelpful    | Slightly Unhelpful    | Neutral   | Slightly Helpful    | Very helpful    | N / A |
| f. error messages                      | Very Unhelpful    | Slightly Unhelpful    | Neutral   | Slightly Helpful    | Very helpful    | N / A |

4. Learning

|                                              | 1              | 2                  | 3       | 4             | 5         | 0     |
|----------------------------------------------|----------------|--------------------|---------|---------------|-----------|-------|
| a. learning to operate the system            | Very Difficult | Slightly Difficult | Neutral | Slightly Easy | Very Easy | N / A |
| b. getting started with the system           | Very Difficult | Slightly Difficult | Neutral | Slightly Easy | Very Easy | N / A |
| c. time to learn the system                  | Very Slow      | Slightly Slow      | Neutral | Slightly Fast | Very Fast | N / A |
| d. remembering rules about entering commands | Very Difficult | Slightly Difficult | Neutral | Slightly Easy | Very Easy | N / A |

|                                                       |                |                     |           |                  |              |       |
|-------------------------------------------------------|----------------|---------------------|-----------|------------------|--------------|-------|
| e. the system functionality and capability            | Not Enough     | Slightly not enough | Neutral   | Slightly Enough  | Enough       | N / A |
| f. tasks can be performed in a straightforward manner | Never          | Almost Never        | Sometimes | Almost Always    | Always       | N / A |
| g. number of steps per task                           | Too Many       | Slightly too many   | Neutral   | About Right      | Just right   | N / A |
| h. keyboard shortcuts                                 | Very Unhelpful | Slightly Unhelpful  | Neutral   | Slightly Helpful | Very helpful | N / A |
| i. steps to complete task follow a logical sequence   | Never          | Almost Never        | Sometimes | Almost Always    | Always       | N / A |

5. System Capabilities

|                                                  |                 |                     |           |               |           |       |
|--------------------------------------------------|-----------------|---------------------|-----------|---------------|-----------|-------|
| a. system speed is                               | Very Slow       | Slightly Slow       | Neutral   | Slightly Fast | Very Fast | N / A |
| b. response time for the operation is            | Very Slow       | Slightly Slow       | Neutral   | Slightly Fast | Very Fast | N / A |
| c. the rate at which information is displayed is | Very Slow       | Slightly Slow       | Neutral   | Slightly Fast | Very Fast | N / A |
| d. system failures occur                         | Very Frequently | Slightly Frequently | Sometimes | Almost Never  | Never     | N / A |

|                                                       |                      |                        |         |                  |           |       |
|-------------------------------------------------------|----------------------|------------------------|---------|------------------|-----------|-------|
| e. novices can accomplish tasks after proper training | With Much Difficulty | With Slight Difficulty | Neutral | With Slight Ease | With Ease | N / A |
| f. correcting mistakes is                             | Very Difficult       | Slightly Difficult     | Neutral | Slightly Easy    | Very Easy | N / A |

6. User Comments

|                                               | 1                 | 2        | 3       | 4     | 5              | 0     |
|-----------------------------------------------|-------------------|----------|---------|-------|----------------|-------|
| a. I like using the interface for this system | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree | N / A |
| 6b. the interface of this system is pleasant  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree | N / A |
| 6c. overall, I am satisfied with this system  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree | N / A |
| 6d. it was easy to learn this system          | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree | N / A |
| 6e. I feel comfortable using this system      | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree | N / A |

List the three most negative aspects of using this system.

- a.
- b.
- c.

List the three most positive aspects of using this system.

- a.
- b.
- c.

Are there any parts of the system, which you found confusing or difficult to fully understand?

Were there any aspects of the system, which you found particularly irritating although they did not cause major problems?

What are the most common mistakes you made when using the system?

What changes would you make to make it better from the user's point of view?

Is there anything else about the system you would like to add?

Please write user's comments to the task on the back of this page.

## **G. TASK INSTRUCTIONS**

Each participant was provided a task sheet with instructions for completing the task. Task 1a was the training task without the SES. Task 1b was the training task with the SES. Task 2e was the evaluation task without the SES. Task 2f was the evaluation task with the SES. The following pages contain those four instruction sheets.

## **G.1**

### **TASK 1A**

You are the supervisor of a mobile robot. The robot has sent you a distress signal that it is stranded on the third floor of the engineering building. You open the user interface and attempt to determine where the robot is because the map failed to display the robot's true location.

1. Look at the environment map to determine landmarks in the area (i.e. cones, balls, boxes, walls, rooms, etc.)
  2. Look at the sonar display to determine obstacles around the robot.
  3. Look at the laser display to determine obstacles around the robot
  4. Look at the compass display to determine which direction the robot is facing.
  5. On the camera display, tilt the camera up and down, move the camera left and right, and zoom the camera in and out to see the robot's surroundings.
  6. Mark the location of the robot (including the direction the robot is facing) on the map you have been provided.
  7. Mark the location of the orange, yellow, pink and green cones around the robot by writing, 'O', 'Y', 'G' and 'P', respectively.
  8. Write down how you would propose to return the robot to the home position shown on the map. Use directions such as 'turn left', or 'go forward'. It is not necessary to give metric information.
- 
-

## G.2

### TASK 1B

You are the supervisor of a mobile robot. The robot has sent you a distress signal that it is stranded on the third floor of the engineering building. You open the user interface and attempt to determine where the robot is because the map failed to display the robot's true location.

1. Look at the environment map to determine landmarks in the area (i.e. cones, balls, boxes, walls, rooms, etc.)
2. Look at the sonar display to determine obstacles around the robot.
3. Look at the laser display to determine obstacles around the robot
4. Look at the compass display to determine which direction the robot is facing.
5. On the camera display, use the scroll bars to tilt the camera up and down, move the camera left and right, and use the button to zoom the camera in and out to see the robot's surroundings.
6. On the Sensory EgoSphere display, click scan to update the images. Use the left mouse button to zoom-in , use the right mouse button to zoom-out, use ctrl+left button to zoom-in faster, use ctrl+right button to zoom-out faster. Use the arrow keys to turn the sphere right, left, up and down. Use ctrl+arrow keys to turn with bigger steps.
7. Mark the location of the robot (including the direction the robot is facing) on the map you have been provided.

8. Mark the location of the orange, yellow, pink and green cones around the robot by writing, 'O', 'Y', 'G' and 'P', respectively.
9. Write down how you would propose to return the robot to the home position shown on the map. Use directions such as 'turn left', or 'go forward'. It is not necessary to give metric information.

### G.3

#### TASK 2E

You are the commander of a mobile robot. You have been given the mission to use the robot to explore the back hallway of the third floor of the engineering building. In order to complete this mission, you will drive to one given location by using several move to point commands on the interface. You will survey the area and then use the move to object command to drive to the pink box at the end of the hallway. During the survey, record the location of all objects found on the map. The objects are yellow, orange, green and pink cones. Record all objects found while moving as well as those found while stationary. Here is what you need to do to complete the mission.

1. Find the point at the juncture of the hallways on the map (the ball)
2. Click Move to Point on the Landmark Map
3. Click the location to go to on the Map
4. Click OK
5. The robot will begin to move, remember to press **STOP** at any moment if the robot begins to move in the wrong direction.
6. Once the robot signals that it has arrived at the goal point do the following.
  - a. Look at the sonar display to determine obstacles around the robot.
  - b. Look at the laser display to determine obstacles around the robot.
  - c. Look at the compass display to determine which direction the robot is facing.
  - d. Use the camera display to see the robot's surroundings.

7. Record the location of all objects found. Repeat this procedure until you reach the juncture of the hallways.
8. Once you are satisfied that you have located all of the objects, move the robot to the next landmark, the pink box.
9. Click Move To Object
10. Click the icon of the pink box
11. Click OK
12. The robot will begin to move, remember to press the **STOP** at any moment if the robot moves in the wrong direction.
13. Use all of the displays to determine all obstacles and objects around the robot as before.
14. Once you arrive at the pink box, you have completed the mission.

## G.4

### TASK 2F

You are the commander of a mobile robot. You have been given the mission to use the robot to explore the back hallway of the third floor of the engineering building. In order to complete this mission, you will drive to one given location by using several move to point commands on the interface. You will survey the area and then use the move to object command to drive to the pink box at the end of the hallway. During the survey, record the location of all objects found on the map. The objects are yellow, orange, green and pink cones. Record all objects found while moving as well as those found while stationary. Here is what you need to do to complete the mission.

1. Find the point at the juncture of the hallways on the map (the ball)
2. Click Move to Point on the Landmark Map
3. Click the location to go to on the Map
4. Click OK
5. The robot will begin to move, remember to press **STOP** at any moment if the robot begins to move in the wrong direction.
6. Once the robot signals that it has arrived at the goal point do the following.
  - a. Look at the sonar display to determine obstacles around the robot.
  - b. Look at the laser display to determine obstacles around the robot.
  - c. Look at the compass display to determine which direction the robot is facing.
  - d. Use the camera display to see the robot's surroundings.
  - e. Scan the SES to determine obstacles and objects around the robot

7. Record the location of all objects found. Repeat this procedure until you reach the juncture of the hallways.
8. Once you are satisfied that you have located all of the objects, move the robot to the next landmark, the pink box.
9. Click Move To Object
10. Click the icon of the pink box
11. Click OK
12. The robot will begin to move, remember to press the **STOP** at any moment if the robot moves in the wrong direction.
13. Use all of the displays to determine all obstacles and objects around the robot as before.
14. Once you arrive at the pink box, you have completed the mission.

## **H. CONSENT FORMS**

All participants were required to complete the following two consent forms. The forms were explained to them and then witnessed and signed. The first form is the Vanderbilt University consent for research study. The second form is the Vanderbilt University consent for videotaping.

## H.1

### Consent for Research Study

#### Vanderbilt University Institutional Review Board

#### Proposal for Research Using Human Subjects

### Consent for Research Study

This consent form applies to: adults

(e.g. adults, child 7-12 years, parent, legal representative, normal volunteer, etc.)

Name of subject \_\_\_\_\_ Age \_\_\_\_\_

**The following information is provided to inform you about the research project and your participation in the study. Please read this form carefully. Please feel free to ask any questions you may have about this study and the information given below. You will be given an opportunity to ask questions, and your questions will be answered. You will be given a copy of this consent form.**

**1. Purpose of the study.**

The purpose of this study is to evaluate a graphical user interface used to remotely control mobile robots. A graphical user interface is a computer program that lets you give commands to a robot using icons and windows. Typically this is done with a mouse.

You will learn to remotely control the robot and record information about the robot's surroundings. During this process, the researcher will study human-robot interaction for different interfaces. In this study, you will be videotaped and asked to complete questionnaires about your experience.

2. **Description of the procedures to be followed and approximate duration of the study.** (Included is a statement of the procedures that will be done solely for research purposes and those that are considered routine treatment. Also included is information about the costs, if any, of the procedures.)

As a participant in this study you will be asked to:

- complete anonymous questionnaires,
- operate a mobile robot using a graphical user interface from a computer,
- view a videotape of your robot operation, and
- be interviewed about your experience

There will be one pre-experiment questionnaire that will gather information about your level of experience with computers and robotics. The post-task questionnaire will gather information about your experiences while performing the tasks operating the mobile robot. The post-experiment questionnaire will gather information about the study in general, such as your reaction to the user interface. During the operation of the mobile robot, you will use the interface to give the robot commands. At the completion of the task, you will view videotape of your task. Finally, you will be interviewed about your reactions and thoughts while viewing the tape.

The study will last approximately two months and you will be asked to commit to two days. The commitment will be two 75-minute sessions over a three-week period. The schedule for the two days is given below,

| DAY 1                                     |            |
|-------------------------------------------|------------|
| Orientation/Welcome                       | 5 minutes  |
| Consent form/Pre-Experiment Questionnaire | 5 minutes  |
| Training                                  | 20 minutes |
| Training task                             | 30 minutes |
| Video Review                              | 10 minutes |
| Post-Task Questionnaire                   | 5 minutes  |
| Total time Commitment                     | 75 minutes |

| DAY 2                         |            |
|-------------------------------|------------|
| Welcome                       | 2 minutes  |
| Task 1 (run twice)            | 20 minutes |
| Video Review                  | 10 minutes |
| Post-Task Questionnaire       | 5 minutes  |
| Task 2 (run twice)            | 20 minutes |
| Video Review                  | 10 minutes |
| Post-Task Questionnaire       | 5 minutes  |
| Post-Experiment Questionnaire | 3 minutes  |
| Total Time Commitment         | 75 minutes |

There will be no cost for any of this procedure.

**3. Description of the discomforts, inconveniences, and/or risks that can be reasonably expected as a result of participation in this study.**

The level of physical and psychological risk is negligible. You will perform no physical labor other than mouse clicks and recording information. You will have an observer present to control the interface evaluation.

**4. Anticipated benefits resulting from this study:**

a) The potential benefits to science and mankind that may result from this study are:

The development of a more effective and efficient robot user interface. This study will provide valuable insight into a new type of user interface. Additionally, because this study is an integral part of my Ph.D. thesis, this study will help me to prove my hypothesis about the new user interface.

b) The potential benefits to you from this study are:

There are no potential benefits for your participation in this study.

5. **Alternative procedures**

The following are alternative procedures or treatments that may be available to you if you choose not to participate in this study:

6. **Contact information:**

If you should have any questions about this research study, please feel free to contact Carlotta Johnson at 615-322-7269 or my Faculty Advisor, Kazuhiko Kawamura at 615-322-2735.

**For additional information about giving consent or your rights as a participant in this study, please feel free to contact the Vanderbilt University Institutional Review Board Office at (615) 322-2918 or toll free at (866-224-8273).**

**Your rights as a volunteer:**

Your participation in this study is voluntary. You may withdraw from this study at any time without prejudicing your standing within Vanderbilt University or any class.

In the event new information becomes available that may affect the risks or benefits associated with this study or your willingness to participate in it, you will be notified so that you can make an informed decision whether or not to continue your participation in this study.

Efforts will be made to keep the personal information in your research record private and confidential but absolute confidentiality cannot be guaranteed. Your personal information may be disclosed if required by law. Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the National Institutes of Health, the Office of Human Research Protections, the National Cancer Institute, Food and Drug Administration, study sponsor, etc.

If any publications result from this research, you will not be identified by name. Sixteen people (16) in total will be asked to participate in this study. All participants will perform the same tasks with two different graphical user interfaces. All collected data will be coded with a participant ID (not name). All data will be kept in a locked area and an average of all data collected will be presented in all publications. If an individual's data is presented in a figure, names or any identifying information will not be included.

**STATEMENT BY PERSON AGREEING TO PARTICIPATE IN THIS STUDY**

**[ ] I have read this consent form. All my questions have been answered, and I freely and voluntarily choose to participate. I understand that I may withdraw at any time.**

**[ ] The material contained in this consent form has been explained to me verbally. All my questions have been answered, and I freely and voluntarily choose to participate. I understand that I may withdraw at any time.**

\_\_\_\_\_

Date

\_\_\_\_\_

Signature of volunteer

Consent obtained by:

\_\_\_\_\_

Signature

\_\_\_\_\_

Printed Name and Title

## H.2

### **Consent for Videotaping**

#### **Vanderbilt University Institutional Review Board**

#### **Proposal for Research Using Human Subjects**

##### Consent for Videotaping

This consent form applies to: adults

(e.g. adults, child 7-12 years, parent, legal representative, normal volunteer, etc.)

Name of subject \_\_\_\_\_ Age \_\_\_\_\_

The following information is provided to inform you about the research project and your participation in the study. Please read this form carefully. Please feel free to ask any questions you may have about this study and the information given below. You will be given an opportunity to ask questions, and your questions will be answered. You will be given a copy of this consent form.

I understand that as a participant in this study, I will be videotaped for the purpose of tracking my movement as well as a means of verifying results from other data collected. I understand that this videotape may be used for presentations to engineering faculty or in robotics conferences. I give my consent for this use of the videotape, I am aware that I may withdraw from the study at any time without penalty.

I further understand that if I have any comments or concerns resulting from my participation in this study, I may contact Carlotta Johnson at 615-322-7269 or the Faculty Advisor, Kazuhiko Kawamura at 615-322-2735.

For additional information about giving consent or your rights as a participant in this study, please feel free to contact the Vanderbilt University Institutional Review Board Office at (615) 322-2918 or toll free at (866-224-8273).

STATEMENT BY PERSON AGREEING TO PARTICIPATE IN THIS STUDY

☐ **I have read this consent form. All my questions have been answered, and I freely and voluntarily choose to participate. I understand that I may withdraw at any time.**

☐ **The material contained in this consent form has been explained to me verbally. All my questions have been answered, and I freely and voluntarily choose to participate. I understand that I may withdraw at any time.**

|       |                        |
|-------|------------------------|
| _____ | _____                  |
| Date  | Signature of volunteer |

Consent obtained by:

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Printed Name and Title

## **I. EXPERIMENTAL RESULTS**

The purpose of this Appendix is to provide more detail on the statistic analysis results discussed in Chapter VII.

,

Table 1 is a statistical analysis of the number of camera clicks for Task A and Task B dependent upon task order.

Table 1. Task A and Task B Camera Clicks

|            | Task A before Task B       |                           | Task B before Task A       |                           |
|------------|----------------------------|---------------------------|----------------------------|---------------------------|
| Click Type | Task A                     | Task B                    | Task A                     | Task B                    |
| Pan        |                            |                           | $z = -1.53,$<br>$p = 0.13$ | $z = 0.28,$<br>$p = 0.40$ |
| Tilt       |                            |                           | $z = -0.22,$<br>$p = 0.83$ | $z = 0.28,$<br>$p = 0.40$ |
| Zoom-In    | $z = -0.67,$<br>$p = 0.50$ | $z = 0.55,$<br>$p = 0.62$ |                            |                           |
| Zoom-Out   | $z = -0.94,$<br>$p = 0.35$ | $z = 0.65,$<br>$p = 0.85$ |                            |                           |
| Reset      |                            |                           | $z = -0.22,$<br>$p = 0.82$ | $z = 0.85,$<br>$p = 0.85$ |
| Total      |                            |                           | $z = -0.65,$<br>$p = 0.51$ | $z = 0.47,$<br>$p = 0.62$ |

Table 2 is a statistical analysis of the number of camera clicks for Task E and Task F dependent upon task order.

Table 2. Tasks E and F Camera Clicks

|            | Task E before Task F |        | Task F before Task E       |                            |
|------------|----------------------|--------|----------------------------|----------------------------|
| Click Type | Task E               | Task F | Task E                     | Task F                     |
| Pan        |                      |        | $z = -0.31,$<br>$p = 0.75$ | $z = -0.94,$<br>$p = 0.34$ |

Table 2, continued

|          |                             |                             |                             |                             |
|----------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Tilt     | $z = -0.73$ ,<br>$p = 0.75$ |                             |                             | $z = -0.83$ ,<br>$P = 0.40$ |
| Zoom-In  | $z = -0.32$ ,<br>$p = 0.75$ |                             |                             | $z = -0.54$ ,<br>$p = 0.58$ |
| Zoom-Out |                             |                             | $z = -0.24$ ,<br>$p = 0.81$ | $z = -0.14$ ,<br>$p = 0.88$ |
| Reset    | $z = -0.54$ ,<br>$p = 0.59$ | $z = -0.21$ ,<br>$z = 0.83$ |                             |                             |
| Total    |                             |                             | $z = -0.31$ ,<br>$p = 0.75$ | $z = -0.73$ ,<br>$p = 0.46$ |

Table 3 is a statistical analysis of the number of map clicks for Task E and Task F dependent upon task order.

Table 3. Tasks E and F Map Clicks

|            | Task E before Task F        |                             | Task F before Task E        |                             |
|------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Click Type | Task E                      | Task F                      | Task E                      | Task F                      |
| Move To    | $z = -0.61$ ,<br>$p = 0.54$ | $z = -0.25$ ,<br>$p = 0.81$ |                             |                             |
| Map        | $a = -0.24$<br>$p = 0.81$   | $z = -0.12$ ,<br>$p = 0.90$ |                             |                             |
| Add Icon   |                             |                             | $z = -0.25$ ,<br>$p = 0.80$ | $z = -0.13$ ,<br>$p = 0.90$ |
| Total      | $z = 0.0$ ,<br>$p = 1.0$    | $z = -0.13$ ,<br>$p = 0.90$ |                             |                             |

Table 4 is a statistical analysis of the number of SES clicks for Task E and Task F dependent upon task order.

Table 4. Task B and Task F SES Clicks

|            | Task B                     |                             | Task F |                             |
|------------|----------------------------|-----------------------------|--------|-----------------------------|
| Click Type | AB                         | BA                          | EF     | FE                          |
| Scan       |                            | $z = -1.18$<br>$p = 0.239$  |        | $z = -0.75$<br>$p = 0.45$   |
| Reset      |                            | $z = -1.27$<br>$p = 0.20$   |        | $z = -0.74$<br>$p = 0.46$   |
| Zoom-in    |                            | $z = -0.44$<br>$p = 0.65$   |        | $z = -1.45$<br>$p = 0.15$   |
| Zoom-out   |                            | $z = -0.53$<br>$p = 0.59$   |        | $z = -2.2$<br>$p = 0.025$   |
| Pan Left   |                            | $z = -0.56$<br>$p = 0.58$   |        | $z = -2.117$<br>$p = 0.034$ |
| Pan Right  |                            | $z = -0.54$<br>$p = 0.59$   |        | $z = -1.94$<br>$p = 0.052$  |
| Tilt Up    |                            | $z = -0.235$<br>$p = 0.814$ |        | $z = -1.53$<br>$p = 0.126$  |
| Tilt Down  | $z = 0.814$<br>$p = 0.841$ |                             |        | $z = -1.29$<br>$p = 0.196$  |
| Total      |                            | $z = -0.313$<br>$p = 0.754$ |        | $z = -1.776$<br>$p = 0.076$ |

Table 5 is a correlation analysis between the number of SES clicks for Task B versus Task F. The remainder of the SES click types did not exhibit any significant correlations.

Table 5. Task B SES Clicks versus Task F SES Clicks Correlation

| Task B     | Task F                              |                                     |                                     |                                   |                                     |
|------------|-------------------------------------|-------------------------------------|-------------------------------------|-----------------------------------|-------------------------------------|
| Click Type | Zoom-in                             | Zoom-out                            | Pan Right                           | Tilt Up                           | Total                               |
| Pan Left   | (+) corr.<br>r = 0.972<br>p = 0.0   | (+) corr.<br>r = 0.774<br>p = 0.024 |                                     | (+) corr.<br>r = 0.968<br>p = 0.0 | (+) corr.<br>r = 0.749<br>p = 0.033 |
| Total      | (+) corr.<br>r = 0.737<br>p = 0.037 | (+) corr.<br>r = 0.755<br>p = 0.03  | (+) corr.<br>r = 0.840<br>p = 0.009 |                                   | (+) corr.<br>r = 0.76<br>p = 0.029  |

Table 6 is a correlation analysis of the number of SES clicks versus the number of camera clicks for Tasks E and F. There were no significant correlations for the remainder the camera click types.

Table 6. SES Clicks versus Camera Clicks Correlation

| Camera Click Type | Task B                                                      | Task F                                                      |
|-------------------|-------------------------------------------------------------|-------------------------------------------------------------|
| Zoom-out          | (-) corr. with SES zoom-out clicks<br>r = -0.771, p = 0.042 |                                                             |
| Reset             |                                                             | (+) corr. with SES tilt down clicks.<br>r = .73, p = 0.017) |

Table 7 is a statistical analysis of the Task A and Task B task scores dependent upon task order.

Table 7. Tasks A and B Scores Statistics

|                   | Task A before Task B       |                           | Task B before Task A    |                            |
|-------------------|----------------------------|---------------------------|-------------------------|----------------------------|
| Sub-Score         | Task A                     | Task B                    | Task A                  | Task B                     |
| Robot Placement   |                            |                           | $z = 0.0,$<br>$p = 1.0$ | $z = -1.67,$<br>$p = 0.09$ |
| Robot Orientation |                            |                           | $z = 0.0,$<br>$p = 1.0$ | $z = -0.73,$<br>$p = 0.46$ |
| Cone Placement    | $z = 0.0,$<br>$p = 1.0$    |                           |                         | $z = -1.5,$<br>$p = 0.13$  |
| Cone Color        | $z = -1.29,$<br>$p = 0.19$ | $z = 0.0,$<br>$p = 1.0$   |                         |                            |
| Driving Direction |                            | $z = 0.52,$<br>$p = 0.85$ |                         |                            |
| Overall Score     | $z = 0.0,$<br>$p = 1.0$    |                           |                         | $z = -0.9,$<br>$p = 0.36$  |

Table 8 is a statistical analysis of the task E and F task scores dependent upon task order.

Table 8. Tasks E and F Scores Statistics

|                | Task E before Task F |        | Task F before Task E       |                            |
|----------------|----------------------|--------|----------------------------|----------------------------|
| Sub-Score      | Task E               | Task F | Task E                     | Task F                     |
| Cone Placement |                      |        | $z = -0.86,$<br>$p = 0.38$ | $z = -0.95,$<br>$p = 0.33$ |
| Cone Color     |                      |        | $z = -0.51,$<br>$p = 0.60$ | $z = -0.11,$<br>$p = 0.91$ |
| Overall Score  |                      |        | $z = -0.84,$<br>$p = 0.39$ | $z = -0.52,$<br>$p = 0.59$ |

Table 9 is a correlation analysis of the number of camera clicks and the task score. The remainder of the camera click types did not exhibit any significant correlations with the task score.

Table 9: Task Score versus Camera Clicks Correlation

| Camera Click Type | Task A                                                     | Task B                                                     | Task E                                                 | Task F                                                |
|-------------------|------------------------------------------------------------|------------------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------|
| Pan               | (-) corr. driving direction<br>$r = -0.859$<br>$p = 0.029$ |                                                            |                                                        |                                                       |
| Tilt              |                                                            | (-) corr. driving direction<br>$r = -0.987$<br>$p = 0.0$   |                                                        |                                                       |
| Zoom-out          |                                                            | (-) corr. robot placement<br>$r = -0.764$<br>$p = 0.046$   | (-) corr. overall score<br>$r = -0.748$<br>$p = 0.013$ |                                                       |
| Reset             | (-) corr. driving direction<br>$r = -0.959$<br>$p = 0.002$ | (-) corr. driving direction<br>$r = -0.956$<br>$p = 0.003$ |                                                        | (+) corr. cone placement<br>$r = 0.717$<br>$p = 0.02$ |
| Total             | (-) corr. driving direction<br>$r = -0.826$<br>$p = 0.043$ |                                                            |                                                        |                                                       |

Table 10 is a correlation analysis of the number of SES clicks and the task score for tasks A and B. The remainder of the SES click types did not exhibit any significant correlations.

Table 10. Task Score versus SES Clicks Correlation

| Click Type | Task B                                                 | Task F                                          |
|------------|--------------------------------------------------------|-------------------------------------------------|
| Pan left   | (-) corr. cone color<br>$r = -0.679, p = 0.064$        |                                                 |
| Pan Right  | (-) corr. robot orientation<br>$r = -0.917, p = 0.001$ |                                                 |
| Tilt Up    |                                                        | (-) corr. cone color<br>$r = -0.675, p = 0.032$ |
| Total      | (-) corr. robot orientation<br>$r = -0.810, p = 0.015$ |                                                 |
| Total      | (-) corr. cone color<br>$r = -0.844, p = 0.008$        |                                                 |

Table 11 is a statistical analysis of the task completion time dependent upon task presentation for Tasks A and B.

Table 11. Task A and B Completion Times

| Sub-Task        | Task A before Task B |                            | Task B before Task A        |        |
|-----------------|----------------------|----------------------------|-----------------------------|--------|
|                 | Task A               | Task B                     | Task A                      | Task B |
| Find the Robot  |                      | $z = -1.65,$<br>$p = 0.09$ | $z = -1.93,$<br>$p = 0.053$ |        |
| Drive the Robot |                      | $z = -0.14,$<br>$p = 0.88$ | $z = -1.04,$<br>$p = 0.29$  |        |
| Overall Time    |                      | $z = -1.04,$<br>$p = 0.29$ | $z = -1.64,$<br>$p = 0.10$  |        |

Table 12 is a correlation analysis between the number of camera clicks and the task completion time. There were no significant correlations for Task E. The remainder of the Camera click types did not exhibit any significant correlations.

Table 12: Task Completion Time versus Camera Clicks Correlation

| Camera Click Type | Task A                                           | Task B                                                         | Task E | Task F |
|-------------------|--------------------------------------------------|----------------------------------------------------------------|--------|--------|
| Pan               |                                                  | (+) corr. driving directions time<br>$r = 0.84$<br>$p = 0.016$ |        |        |
| Tilt              |                                                  | (+) corr. find the robot time<br>$r = 0.801$<br>$p = 0.03$     |        |        |
| Tilt              |                                                  | (+) corr. total time<br>$r = 0.79$<br>$p = 0.033$              |        |        |
| Zoom-Out          | (+) corr. total time<br>$r = 0.83$<br>$p = 0.04$ |                                                                |        |        |
| Reset             |                                                  | (+) corr. find the robot time<br>$r = 0.94$<br>$p = 0.001$     |        |        |
| Reset             |                                                  | (+) corr. total time<br>$r = 0.94$<br>$p = 0.002$              |        |        |

Table 12, continued

|       |  |                                                                  |  |                                                       |
|-------|--|------------------------------------------------------------------|--|-------------------------------------------------------|
| Total |  | (+) corr. drive<br>directions time<br>$r = 0.861$<br>$p = 0.013$ |  | (+) corr. total<br>time<br>$r = 0.713$<br>$p = 0.021$ |
|-------|--|------------------------------------------------------------------|--|-------------------------------------------------------|

Table 13 is a correlation analysis between the number of SES clicks and the task completion time for Tasks B and F. There were no significant correlations for Task F. There were no significant correlations for the remainder of the SES click types.

Table 13. Task Completion Time versus SES Clicks Correlation

| SES click type | Task B                                                   | Task F |
|----------------|----------------------------------------------------------|--------|
| Zoom-out       | (-) corr. find the robot time<br>$r = -0.722, p = 0.043$ |        |
| Zoom-out       | (-) corr. total task time<br>$r = -0.717, p = 0.045$     |        |

Table 14 is a correlation analysis between the task score and the task completion time. There were not significant correlations between Tasks E and F. The remainder of the score sub-tasks did not exhibit any significant correlations.

Table 14. Task Completion Time versus Task Score Correlation

| Score Sub-Task     | Task A | Task B                                                    | Task E | Task F |
|--------------------|--------|-----------------------------------------------------------|--------|--------|
| Driving Directions |        | (-) corr. find the robot time<br>$r = -0.99$<br>$p = 0.0$ |        |        |
| Driving Directions |        | (-) corr. total task time<br>$r = -0.99$<br>$p = 0.0$     |        |        |

Table 15 is a statistical analysis of the responses to the MRQ dependent upon task order for Tasks A and B.

Table 15. Tasks A and Task B MRQ

|                     | Task A before Task B        |        | Task B before Task A        |                              |
|---------------------|-----------------------------|--------|-----------------------------|------------------------------|
| Sub-Process         | Task A                      | Task B | Task A                      | Task B                       |
| Manual              |                             |        | $z = -0.31$ ,<br>$p = 0.75$ | $z = -0.77$ ,<br>$p = 0.439$ |
| Short-term memory   | $z = -1.32$ ,<br>$p = 0.18$ |        |                             | $z = 0.0$ ,<br>$p = 1.0$     |
| Spatial attentive   | $z = -0.89$ ,<br>$p = 0.36$ |        |                             | $z = -1.3$ ,<br>$p = 0.19$   |
| Spatial categorical | $z = -0.93$ ,<br>$p = 0.35$ |        |                             | $z = -0.19$ ,<br>$p = 0.84$  |
| Spatial emergent    |                             |        | $z = -1.1$ ,<br>$p = 0.26$  | $z = -0.17$ ,<br>$p = 0.86$  |

Table 15, continued

|                      |                             |                             |                             |                             |
|----------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Spatial positional   | $z = -0.15$ ,<br>$p = 0.87$ |                             |                             | $z = -0.34$<br>$p = 0.73$   |
| Spatial quantitative |                             | $z = -0.19$ ,<br>$p = 0.84$ | $z = -0.66$ ,<br>$p = 0.50$ |                             |
| Visual lexical       | $z = -0.66$ ,<br>$p = 0.50$ | $z = -1.14$ ,<br>$p = 0.25$ |                             |                             |
| Visual temporal      | $z = -1.08$ ,<br>$p = 0.27$ |                             |                             | $z = -0.68$ ,<br>$p = 0.49$ |
| Overall Ratings      | $z = -0.73$ ,<br>$p = 0.46$ |                             |                             | $z = -0.30$ ,<br>$p = 0.76$ |

Table 16 is a statistical analysis of the responses to the MRQ dependent upon task order for Tasks E and F.

Table 16. Tasks E and Task F MRQ

| Sub-Process         | Task E before Task F        |                             | Task F before Task E        |                             |
|---------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|                     | Task E                      | Task F                      | Task E                      | Task F                      |
| Manual              |                             |                             |                             |                             |
| Short-term memory   | $z = -0.51$ ,<br>$p = 0.57$ |                             |                             | $z = -0.51$ ,<br>$p = 0.60$ |
| Spatial attentive   |                             | $z = -0.80$ ,<br>$p = 0.41$ | $z = -1.06$ ,<br>$p = 0.28$ |                             |
| Spatial categorical | $z = -0.55$ ,<br>$p = 0.57$ | $z = -0.11$ ,<br>$p = 0.91$ |                             |                             |
| Spatial emergent    | $z = -0.43$<br>$p = 0.66$   |                             |                             | $z = -0.95$ ,<br>$p = 0.33$ |
| Spatial positional  | $z = -0.77$ ,<br>$p = 0.43$ | $z = -1.10$ ,<br>$p = 0.26$ |                             |                             |

Table 16, continued

|                      |                           |                             |                             |  |
|----------------------|---------------------------|-----------------------------|-----------------------------|--|
| Spatial quantitative | $z = -1.75$<br>$p = 0.08$ | $z = -0.21$ ,<br>$p = 0.82$ |                             |  |
| Visual lexical       | $z = -0.65$<br>$p = 0.51$ | $z = -0.54$ ,<br>$p = 0.58$ |                             |  |
| Visual temporal      |                           | $z = -0.11$ ,<br>$p = 0.90$ | $z = -0.55$ ,<br>$p = 0.57$ |  |
| Overall Ratings      | $z = -0.73$<br>$p = 0.46$ | $z = -0.53$<br>$p = 0.59$   |                             |  |

Table 17 is a correlation analysis between the responses to the MRQ and the number of camera clicks. Task A did not exhibit any significant correlations. The remainder of the click types did not exhibit any significant correlations.

Table 17: MRQ versus Camera Clicks

| Camera Click Type | Task A | Task B                                                  | Task E | Task F                                                     |
|-------------------|--------|---------------------------------------------------------|--------|------------------------------------------------------------|
| Pan               |        |                                                         |        | (-) corr. visual temporal<br>$r = -0.818$<br>$p = 0.007$   |
| Pan               |        |                                                         |        | (-) corr. spatial attentive<br>$r = -0.679$<br>$p = 0.031$ |
| Tilt              |        | (-) corr. visual temporal<br>$r = -0.835$<br>$p = .019$ |        |                                                            |

Table 17, continued

|          |  |                                                                  |                                                             |                                                                  |
|----------|--|------------------------------------------------------------------|-------------------------------------------------------------|------------------------------------------------------------------|
| Zoom-In  |  | (-) corr. spatial<br>quantitative<br>$r = -0.862$<br>$p = 0.013$ | (-) corr. overall<br>average<br>$r = -0.668$<br>$p = 0.035$ | (-) corr. spatial<br>quantitative<br>$r = -0.664$<br>$p = 0.036$ |
| Zoom-Out |  | (-) corr. spatial<br>quantitative<br>$r = -0.85$<br>$p = 0.014$  |                                                             | (+) corr.<br>manual<br>processes<br>$r = 0.739$<br>$p = 0.015$   |
| Reset    |  | (-) corr. spatial<br>quantitative<br>$r = -0.95$<br>$p = 0.001$  |                                                             | (-) corr. visual<br>temporal<br>$r = -0.693$<br>$p = 0.039$      |
| Total    |  |                                                                  | (-) corr. manual<br>process<br>$r = -0.647$<br>$p = 0.043$  | (-) corr. visual<br>temporal<br>$r = -0.799$<br>$p = 0.01$       |
| Total    |  |                                                                  | (-) corr. visual<br>temporal<br>$r = -0.631$<br>$p = 0.05$  | (-) corr. spatial<br>attentive<br>$r = -0.667$<br>$p = 0.035$    |

Table 18 is a correlation analysis between the responses to the MRQ and the number of map clicks for Tasks E and F. Task F did not exhibit any significant correlations between the MRQ and the number of map clicks.

Table 18: MRQ versus Map Clicks Correlation

| Map Click<br>Type | Task E                                     | Task F |
|-------------------|--------------------------------------------|--------|
| Move to<br>Point  | (+) corr. manual<br>$r = 0.678, p = 0.045$ |        |

|               |                                                                                                  |  |
|---------------|--------------------------------------------------------------------------------------------------|--|
| Move to Point | (+) corr. visual temporal<br>$r = 0.75, p = 0.02$                                                |  |
| Map           | (+) corr. manual<br>$r = 0.704, p = 0.034$                                                       |  |
| Map           | (+) corr. visual temporal<br>$r = 0.691, p = 0.039$                                              |  |
| Add Icon      | (+) corr. visual temporal<br>$r = 0.692, p = 0.039$                                              |  |
| Total         | (+) corr. manual<br>$r = 0.68, p = 0.044$<br>(+) corr. visual temporal<br>$r = 0.694, p = 0.038$ |  |

Table 19 is a correlation analysis between the responses to the MRQ and the number of SES clicks for Tasks B and F. The remainder of the SES click types did not exhibit any significant correlations.

Table 19: MRQ versus SES Clicks Correlation

| Click Type | Task B                                                   | Task F                                                    |
|------------|----------------------------------------------------------|-----------------------------------------------------------|
| Zoom-in    |                                                          | (-) corr. visual temporal<br>$r = -0.653, p = 0.041$      |
| Zoom-out   | (+) corr. spatial quantitative<br>$r = 0.861, p = 0.006$ | (+) corr. spatial quantitative<br>$r = 0.69, p = 0.027$   |
| Pan left   | (+) corr. spatial positional<br>$r = 0.772, p = 0.025$   | (+) corr. spatial quantitative<br>$r = 0.717, p = 0.02$   |
| Pan right  | (+) corr. overall ratings<br>$r = 0.764, p = 0.027$      | (+) corr. spatial quantitative<br>$r = 0.878, p = 0.0001$ |

Table 19, continued

|           |  |                                                          |
|-----------|--|----------------------------------------------------------|
| Tilt down |  | (-) corr. spatial emergent<br>$r = -0.687, p = 0.028$    |
| Total     |  | (+) corr. spatial quantitative<br>$r = 0.807, p = 0.005$ |

Table 20 is a correlation analysis between the responses to the MRQ and the task score. The remainder of the MRQ ratings did not exhibit any significant correlations.

Table 20. MRQ versus Task Score Correlation

|                      | Task A                                                          | Task B                                                          | Task E | Task F                                                |
|----------------------|-----------------------------------------------------------------|-----------------------------------------------------------------|--------|-------------------------------------------------------|
| Spatial quantitative |                                                                 | (+) corr. driving direction. score<br>$r = 0.88$<br>$p = 0.009$ |        |                                                       |
| Visual lexical       | (-) corr. overall score<br>$r = -0.74$<br>$p = 0.04$            |                                                                 |        |                                                       |
| Visual temporal      |                                                                 | (+) corr. driving directions score<br>$r = 0.76$<br>$p = 0.046$ |        |                                                       |
| Overall Ratings      | (-) corr. driving directions score<br>$r = -0.72$<br>$p = 0.04$ |                                                                 |        | (-) corr. overall score<br>$r = -0.77$<br>$p = 0.009$ |

Table 21 is a correlation analysis between the responses to the MRQ and the task completion time. The remainder of the MRQ ratings did not exhibit significant correlations. There were no significant correlations for Tasks E and F.

Table 21: MRQ versus Time Correlation

| Questions            | Task A                                                              | Task B                                                                | Task E | Task F |
|----------------------|---------------------------------------------------------------------|-----------------------------------------------------------------------|--------|--------|
| Spatial emergent     | (-) corr. with the overall task time<br>$r = -0.754$<br>$p = 0.031$ |                                                                       |        |        |
| Spatial quantitative |                                                                     | (-) corr. with the find the robot time<br>$r = -0.893$<br>$p = 0.003$ |        |        |
| Spatial quantitative |                                                                     | (-) corr. with the overall task time<br>$r = -0.873$<br>$p = 0.005$   |        |        |

Table 22 is a statistical analysis for the responses to the NASA-TLX dependent upon the task order for Tasks A and B. The remainder of the NASA-TLX ratings did not exhibit significant correlations. The overall rating for the NASA-TLX was an average of the individual sub-ratings.

Table 22. Tasks A and Task B NASA-TLX

|                          | Task A before Task B         |                             | Task B before Task A          |                            |
|--------------------------|------------------------------|-----------------------------|-------------------------------|----------------------------|
| Sub-Rating               | Task A                       | Task B                      | Task A                        | Task B                     |
| Necessary thinking       | $z = -0.75$ ,<br>$p = 0.45$  |                             |                               | $z = -0.75$<br>$p = 0.46$  |
| Task difficulty          | $z = -0.74$ ,<br>$p = 0.45$  | $z = -0.44$ ,<br>$p = 0.65$ |                               |                            |
| Physical demand          | $z = -0.77$ ,<br>$p = 0.43$  |                             |                               | $z = 0.0$<br>$p = 1.0$     |
| Time required            | $z = -0.44$ ,<br>$p = 0.65$  | $z = -0.74$ ,<br>$p = 0.45$ |                               |                            |
| Time pressure            | $z = -1.11$ ,<br>$p = 0.26$  |                             |                               | $z = 0.0$<br>$p = 1.0$     |
| Goal achievement         | $z = -0.31$ ,<br>$p = 0.75$  |                             |                               | $z = -0.29$<br>$p = 0.76$  |
| Performance satisfaction | $p = -0.76$ ,<br>$p = 0.44$  | $z = -0.76$<br>$p = 0.44$   |                               |                            |
| Mental effort            | $z = -2.23$ ,<br>$p = 0.025$ | $z = -0.14$<br>$p = 0.88$   |                               |                            |
| Physical effort          |                              |                             | $z = -0.195$ ,<br>$p = 0.844$ | $z = -1.17$<br>$p = 0.24$  |
| Frustration level        | $z = -1.68$<br>$p = 0.09$    |                             |                               | $z = -1.53$<br>$p = 0.124$ |
| Stress level             | $z = 0.24$<br>$p = 0.24$     |                             |                               | $z = -0.34$<br>$p = 0.73$  |
| Overall Rating           | $z = -1.63$ ,<br>$p = 0.10$  |                             |                               | $z = -0.74$<br>$p = 0.45$  |

Table 23 is a statistical analysis for the responses to the NASA-TLX dependent upon the task order for Tasks E and F. The remainder of the NASA-TLX ratings did not exhibit significant correlations.

Table 23. Tasks E and Task F NASA-TLX

|                          | Task E before Task F         |                              | Task F before Task E         |                             |
|--------------------------|------------------------------|------------------------------|------------------------------|-----------------------------|
| Sub-Rating               | Task E                       | Task F                       | Task E                       | Task F                      |
| Necessary thinking       | $z = -1.35$ ,<br>$p = 0.174$ | $z = -0.1$ ,<br>$p = 0.917$  |                              |                             |
| Task difficulty          | $z = -1.14$ ,<br>$p = 0.25$  |                              |                              | $z = -0.73$ ,<br>$p = 0.47$ |
| Physical demand          | $z = 0.0$ ,<br>$p = 1.0$     |                              |                              | $z = -1.52$ ,<br>$p = 0.12$ |
| Time required            | $z = -1.57$ ,<br>$p = 0.12$  | $z = -0.83$ ,<br>$p = 0.40$  |                              |                             |
| Time pressure            |                              |                              | $z = -0.31$ ,<br>$p = 0.75$  | $z = -0.10$ ,<br>$p = 0.91$ |
| Goal achievement         |                              | $p = -0.73$ ,<br>$p = 0.464$ | $z = -0.31$ ,<br>$p = 0.754$ |                             |
| Performance satisfaction |                              | $z = -0.62$ ,<br>$p = 0.52$  | $z = -0.10$ ,<br>$p = 0.91$  |                             |
| Mental effort            | $z = -0.10$ ,<br>$p = 0.91$  | $z = -0.10$ ,<br>$p = 0.91$  |                              |                             |
| Physical effort          |                              | $z = -0.35$ ,<br>$p = 0.72$  | $z = -1.29$ ,<br>$p = 0.19$  |                             |
| Frustration level        | $z = -0.52$ ,<br>$p = 0.60$  |                              |                              | $z = -0.73$ ,<br>$p = 0.46$ |
| Stress level             |                              | $z = -1.15$ ,<br>$p = 0.24$  | $z = -1.39$ ,<br>$p = 0.163$ |                             |
| Overall Rating           | $z = -0.52$ ,<br>$p = 0.6$   | $z = -0.10$ ,<br>$p = 0.91$  |                              |                             |

Table 24 is a correlation analysis between the responses to the NASA-TLX and the number of camera clicks. The remainder of the NASA-TLX ratings did not exhibit

significant correlations. The remainder of the NASA-TLX ratings did not exhibit significant correlations.

Table 24. NASA-TLX versus Camera Clicks Correlation

| Sub-Rating         | Task A                                                     | Task B                                                       | Task E                                                       | Task F                                                       |
|--------------------|------------------------------------------------------------|--------------------------------------------------------------|--------------------------------------------------------------|--------------------------------------------------------------|
| Necessary thinking | (+) corr. with zoom-out clicks<br>$r = 0.88$<br>$p = 0.02$ |                                                              |                                                              |                                                              |
| Time required      | (+) corr. with reset clicks<br>$r = 0.893$<br>$p = 0.02$   | (+) corr with zoom-in clicks<br>$r = 0.861$<br>$p = 0.013$   | (+) corr with zoom-out clicks<br>$r = 0.664$<br>$p = 0.036$  |                                                              |
| Time pressure      |                                                            |                                                              | (+) corr. with zoom-out clicks<br>$r = 0.693$<br>$p = 0.026$ |                                                              |
| Goal achievement   | (-) corr. with reset clicks<br>$r = -0.83$<br>$p = 0.04$   | (-) corr. with zoom-out clicks<br>$r = -0.96$<br>$p = 0.001$ |                                                              | (-) corr. with zoom-in clicks<br>$r = -0.693$<br>$p = 0.026$ |
| Mental effort      |                                                            | (+) corr with zoom-in clicks<br>$r = 0.975$<br>$p = 0.0$     |                                                              |                                                              |
| Physical effort    |                                                            | (+) corr with pan clicks<br>$r = 0.77$<br>$p = 0.04$         |                                                              |                                                              |
| Frustration level  |                                                            | (+) corr. with tilt clicks<br>$r = 0.788$<br>$p = 0.035$     |                                                              |                                                              |

Table 24, continued

|              |  |                                                              |  |  |
|--------------|--|--------------------------------------------------------------|--|--|
| Stress level |  | (+) corr. with<br>reset clicks<br>$r = 0.959$<br>$p = 0.001$ |  |  |
|--------------|--|--------------------------------------------------------------|--|--|

Table 25 is a correlation analysis between the responses to the NASA-TLX and the number of map clicks. The remainder of the NASA-TLX ratings did not exhibit significant correlations.

Table 25. NASA-TLX versus Map Clicks Correlation

| Sub-Rating         | Task E                                                   | Task F                                                      |
|--------------------|----------------------------------------------------------|-------------------------------------------------------------|
| Necessary thinking | (-) corr. with add icon clicks<br>$r = -0.74, p = 0.021$ |                                                             |
| Frustration level  | (-) corr with add icon clicks<br>$r = -0.67, p = 0.05$   |                                                             |
| Overall Rating     | (-) corr. with add icon clicks<br>$r = -0.68, p = 0.04$  | (+) corr. with map clicks<br>$r = 0.67, p = 0.05$           |
| Overall Rating     |                                                          | (+) corr. with add icon<br>clicks<br>$r = 0.691, p = 0.039$ |

Table 26 is a correlation analysis between the responses to the NASA-TLX and the number of SES clicks. The remainder of the NASA-TLX ratings did not exhibit significant correlations.

Table 26. NASA-TLX versus SES Clicks Correlation

| Sub-Rating         | Task B                                                      | Task F                                                       |
|--------------------|-------------------------------------------------------------|--------------------------------------------------------------|
| Necessary thinking | (-) corr. with SES scan clicks<br>$r = -0.78, p = 0.02$     | (-) corr with SES pan right clicks<br>$r = -0.636, p = 0.04$ |
| Task difficulty    | (+) corr. with SES zoom-in clicks<br>$r = 0.71, p = 0.04$   |                                                              |
| Mental effort      | (+) corr. with scan clicks<br>$r = 0.719, p = 0.04$         | (+) corr. with SES scan clicks<br>$r = 0.66, p = 0.04$       |
| Frustration level  | (-) corr. with SES zoom-out clicks<br>$r = -0.72, p = 0.04$ |                                                              |
| Stress level       | (-) corr. with SES zoom-out<br>$r = -0.72, p = 0.05$        |                                                              |

Table 27 is a correlation analysis between the responses to the NASA-TLX and the task score. The remainder of the NASA-TLX ratings did not exhibit significant correlations.

Table 27. NASA-TLX versus Task Score Correlation

| Sub-Rating         | Task A | Task B | Task E                                                      | Task F |
|--------------------|--------|--------|-------------------------------------------------------------|--------|
| Necessary thinking |        |        | (-) corr with cone color score<br>$r = -0.74$<br>$p = 0.01$ |        |
| Task difficulty    |        |        | (-) corr. with overall score<br>$r = -0.64$<br>$p = 0.05$   |        |

Table 27, continued

|                   |                                                                 |                                                                |                                                                   |                                                          |
|-------------------|-----------------------------------------------------------------|----------------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------------------------|
| Time required     |                                                                 | (-) corr. with driving score<br>$r = -0.87$<br>$p = 0.01$      | (-) corr. with overall score<br>$r = -0.85$<br>$p = 0.002$        |                                                          |
| Time pressure     |                                                                 |                                                                | (-) corr. with overall score<br>$r = -0.75$<br>$p = 0.012$        |                                                          |
| Goal achievement  |                                                                 | (+) corr. with robot plcmt score<br>$r = 0.75$<br>$p = 0.03$   | (+) corr. with cone placement score<br>$r = 0.657$<br>$p = 0.039$ |                                                          |
| Mental effort     |                                                                 |                                                                |                                                                   | (+) corr. with overall score<br>$r = 0.66$<br>$p = 0.04$ |
| Frustration level | (-) corr. with driving dir. score<br>$r = -0.86$<br>$p = 0.005$ | (-) corr. with driving dir. score<br>$r = -0.992$<br>$p = 0.0$ | (-) corr. with overall score<br>$r = -0.83$<br>$p = 0.002$        |                                                          |
| Stress level      |                                                                 | (-) corr. with driving dir. score<br>$r = -1.0$<br>$p = 0.0$   |                                                                   |                                                          |

Table 28 is a correlation analysis between the responses to the NASA-TLX and the task completion time. There were no significant correlations for Tasks B, Task E and Task F. The remainder of the NASA-TLX ratings did not exhibit significant correlations.

Table 28: NASA-TLX vs. Time Correlation

|                   | Task A                                                            | Task B | Task E | Task F |
|-------------------|-------------------------------------------------------------------|--------|--------|--------|
| Frustration level | (-) corr. with the driving direction time<br>$r = .808, p = .015$ |        |        |        |

Table 29 is a correlation analysis between the responses to the NASA-TLX and the MRQ. The remainder of the MRQ ratings did not exhibit significant correlations.

Table 29. NASA-TLX versus MRQ Correlation

|                   | Task A                                                          | Task B | Task E                                                         | Task F |
|-------------------|-----------------------------------------------------------------|--------|----------------------------------------------------------------|--------|
| Manual            | (+) corr. with necessary thinking<br>$r = 0.888$<br>$p = 0.003$ |        |                                                                |        |
| Short-term memory | (-) corr. with stress level<br>$r = -0.846$<br>$p = 0.008$      |        | (+) corr. overall workload average<br>$r = 0.63$<br>$p = 0.04$ |        |
| Spatial attentive | (+) corr. with necessary thinking<br>$r = 0.796$<br>$p = 0.018$ |        | (+) corr. overall workload average<br>$r = 0.64$<br>$p = 0.04$ |        |

Table 29, continued

|                      |                                                                 |                                                            |  |                                                                 |
|----------------------|-----------------------------------------------------------------|------------------------------------------------------------|--|-----------------------------------------------------------------|
| Spatial categorical  | (+) corr. with necessary thinking<br>$r = 0.802$<br>$p = 0.017$ |                                                            |  |                                                                 |
| Spatial emergent     | (-) corr. with task difficulty<br>$r = -0.89$<br>$p = 0.003$    |                                                            |  |                                                                 |
| Spatial positional   | (+) corr. with necessary thinking<br>$r = 0.754$<br>$p = 0.031$ |                                                            |  | (-) corr. with necessary thinking<br>$r = -0.75$<br>$p = 0.013$ |
| Spatial positional   |                                                                 |                                                            |  | (-) corr. mental effort<br>$r = -0.75$<br>$p = 0.013$           |
| Spatial quantitative |                                                                 | (-) corr. with time required<br>$r = -0.89$<br>$p = 0.003$ |  | (-) corr. with necessary thinking<br>$r = -0.72$<br>$p = 0.018$ |
| Spatial quantitative |                                                                 | (-) corr. frustration level<br>$r = -0.563$<br>$p = 0.006$ |  | (-) corr. mental effort<br>$r = -0.67$<br>$p = 0.03$            |
| Spatial quantitative |                                                                 | (-) corr. stress level<br>$r = -0.89$<br>$p = 0.003$       |  |                                                                 |

Table 29, continued

|                 |                                                            |  |                                                                                                                             |                                                                                                                               |
|-----------------|------------------------------------------------------------|--|-----------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Visual lexical  | (-) corr. with stress level<br>$r = -0.824$<br>$p = 0.012$ |  |                                                                                                                             |                                                                                                                               |
| Visual temporal | (-) corr. with stress level<br>$r = -0.796$<br>$p = 0.018$ |  |                                                                                                                             |                                                                                                                               |
| Overall Ratings |                                                            |  | (+) corr. with frustration level<br>$r = 0.663$<br>$p = 0.037$<br>(+) corr. with stress level<br>$r = 0.753$<br>$p = 0.012$ | (-) corr. with necessary thinking<br>$r = -0.75$<br>$p = 0.013$<br>(-) corr. with mental effort<br>$r = -0.79$<br>$r = 0.006$ |

Table 30 is a correlation analysis between the spatial reasoning scores and the number of SES clicks. There were no significant correlations with Task B. There were also No other significant correlations with the remainder of the SES click types for Task F.

Table 30: Spatial Reasoning versus SES Clicks

|                   | Task B | Task F                                                |
|-------------------|--------|-------------------------------------------------------|
| Spatial Reasoning |        | (+) corr. with SES scan<br>$r = 0.683, p = 0.037$     |
| Spatial Reasoning |        | (+) corr. with reset clicks<br>$r = 0.894, p = 0.026$ |

Table 31 is a correlation analysis between the spatial reasoning scores and the responses to the MRQ. The remainder of the NASA-TLX ratings did not have significant correlations. The overall NASA-TLX rating was an average of all of the individual ratings.

Table 31. Spatial Reasoning versus NASA-TLX Correlation

| Sub-Rating         | Task A                                                    | Task F                                                     |
|--------------------|-----------------------------------------------------------|------------------------------------------------------------|
| Necessary thinking |                                                           | (-) corr. with spatial reasoning<br>$r = -0.67, p = 0.03$  |
| Task difficulty    |                                                           | (-) corr. with spatial reasoning<br>$r = -0.67, p = 0.031$ |
| Frustration level  | (+) corr. with spatial reasoning<br>$r = 0.81, p = 0.014$ | (-) corr. with spatial reasoning<br>$r = -0.71, p = 0.02$  |
| Overall Rating     |                                                           | (-) corr. with spatial reasoning<br>$r = -0.91, p = 0$     |

Table 32 is a statistical analysis of the responses to task specific ratings dependent upon the task order for Tasks A and B. The remainder of the task questionnaires did not exhibit significant correlations.

Table 32. Task A and Task B Task Specific Ratings

|                   | Task A before Task B       |                            | Task B before Task A |        |
|-------------------|----------------------------|----------------------------|----------------------|--------|
| Questions         | Task A                     | Task B                     | Task A               | Task B |
| Difficulty level  | $z = -0.60$<br>$p = 0.54$  | $z = 0.0$<br>$p = 1.0$     |                      |        |
| Clarity level     | $z = -1.54$<br>$p = 0.121$ | $z = -0.89$<br>$p = 0.371$ |                      |        |
| Stimulation level | $z = -2.39$<br>$p = 0.016$ | $z = -0.46$<br>$p = 0.64$  |                      |        |
| Frustration level | $z = -1.91$<br>$p = 0.05$  | $z = -0.71$<br>$p = 0.47$  |                      |        |

Table 33 is a statistical analysis of the responses to the general questions of the post-task questionnaire dependent upon the task order for Tasks A and B.

Table 33. Task A and Task B General Questions

|                                       | Task A before Task B       |                           | Task B before Task A |                           |
|---------------------------------------|----------------------------|---------------------------|----------------------|---------------------------|
| Questions                             | Task A                     | Task B                    | Task A               | Task B                    |
| Understanding data and interpretation | $z = -0.34$<br>$p = 0.73$  |                           |                      | $z = -1.41$<br>$p = 0.15$ |
| Error correction capabilities         | $z = -1.91$<br>$p = 0.05$  | $z = -0.28$<br>$p = 0.77$ |                      |                           |
| System Control                        | $z = -0.47$<br>$p = 0.638$ | $z = -0.28$<br>$p = 0.77$ |                      |                           |
| Ability to complete tasks             | $z = -1.23$<br>$p = 0.21$  |                           |                      | $z = -0.53$<br>$p = 0.59$ |

Table 34 is a statistical analysis of the responses to the system capability ratings of the post-task questionnaire dependent upon the task order for Tasks A and B.

Table 34. Task A and Task B System Capability Ratings

| Questions                        | Task A before Task B      |                           | Task B before Task A      |                           |
|----------------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
|                                  | Task A                    | Task B                    | Task A                    | Task B                    |
| System speed                     |                           | $z = -0.26$<br>$p = 0.79$ | $z = -0.30$<br>$p = 0.76$ |                           |
| Power level                      |                           | $z = -0.68$<br>$p = 0.49$ | $z = -0.49$<br>$p = 0.62$ |                           |
| Rigidity                         |                           | $z = -1.45$<br>$p = 0.14$ | $z = -0.34$<br>$p = 0.73$ |                           |
| Ease of task completion          | $z = -0.34$<br>$p = 0.73$ | $z = -0.13$<br>$p = 0.89$ |                           |                           |
| Time constraints                 | $z = -0.95$<br>$p = 0.33$ |                           |                           | $z = -0.53$<br>$p = 0.59$ |
| Support information satisfaction | $z = -0.68$<br>$p = 0.49$ | $z = -0.54$<br>$p = 0.58$ |                           |                           |

Table 35 is a statistical analysis of the responses to the task specific ratings of the post-task questionnaire dependent upon the task order for Tasks E and F.

Table 35. Task E and Task F Task Specific Ratings

| Questions        | Task E before Task F     |        | Task F before Task E      |        |
|------------------|--------------------------|--------|---------------------------|--------|
|                  | Task E                   | Task F | Task E                    | Task F |
| Difficulty level | $z = 0.00$<br>$p = 1.00$ |        | $z = -0.79$<br>$p = 0.43$ |        |

Table 35, continued

|                   |  |                           |                           |  |
|-------------------|--|---------------------------|---------------------------|--|
| Clarity level     |  | $z = -1.00$<br>$p = 0.32$ | $z = -1.67$<br>$p = 0.10$ |  |
| Stimulation level |  | $z = -1.25$<br>$p = 0.21$ | $z = -0.95$<br>$p = 0.34$ |  |
| Frustration level |  | $z = -1.08$<br>$p = 0.28$ | $z = -0.34$<br>$p = 0.74$ |  |

Table 36 is a statistical analysis of the responses to the general questions of the post-task questionnaire dependent upon the task order for Tasks E and F.

Table 36. Task E and Task F General Questions

| Questions                             | Task E before Task F      |                           | Task F before Task E      |                           |
|---------------------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
|                                       | Task E                    | Task F                    | Task E                    | Task F                    |
| Understanding data and interpretation | $z = -0.45$<br>$p = 0.65$ |                           |                           | $z = -0.52$<br>$p = 0.61$ |
| Error correction capabilities         |                           | $z = -0.52$<br>$p = 0.61$ | $z = -1.34$<br>$p = 0.18$ |                           |
| System Control                        |                           |                           | $z = -0.83$<br>$p = 0.41$ | $z = -0.95$<br>$p = 0.34$ |
| Ability to complete tasks             |                           |                           | $z = -1.23$<br>$p = 0.22$ | $z = -1.80$<br>$p = 0.07$ |

Table 37 is a statistical analysis of the responses to the system capability ratings of the post-task questionnaire dependent upon the task order for Tasks E and F.

Table 37. Task E and Task F System Capability Ratings

|                                  | Task E before Task F  |                       | Task F before Task E  |                       |
|----------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Questions                        | Task E                |                       | Task E                | Task F                |
| System speed                     |                       |                       | z = -0.45<br>p = 0.66 | z = -0.56<br>p = 0.58 |
| Power level                      |                       | z = -0.11<br>p = 0.91 | z = -0.57<br>p = 0.57 |                       |
| Rigidity                         | z = -1.09<br>p = 0.28 | z = -0.88<br>p = 0.38 |                       |                       |
| Ease of task completion          |                       | z = -0.12<br>p = 0.91 | z = -1.34<br>p = 0.18 |                       |
| Time constraints                 |                       |                       | z = -1.47<br>p = 0.14 | z = -0.88<br>p = 0.38 |
| Support information satisfaction |                       | z = -0.45<br>p = 0.65 | z = -0.12<br>p = 0.91 |                       |

## BIBLIOGRAPHY

- [Abidi et al., 1992] Abidi, M. A., and R.C. Gonzalez, *Data Fusion In Robotics And Machine Intelligence*, pp. 1 – 108, Academic Press, New York, 1992.
- [Adams, 1995] Adams, J. A., Human Management of a Hierarchical System of the Control of Multiple Mobile Robots, Ph.D. Thesis, Computer and Information Science, University of Pennsylvania, Philadelphia, 1995.
- [Albus, 1991] Albus, J. A., Outline for a theory of Intelligence, *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3), pp. 473-509, 1991.
- [Albus, 1996] Albus, J. A., The Engineering of Mind, in Proceedings of the 4<sup>th</sup> *International Conference on Simulation of Adaptive Behavior: From Animals to Animats 4*, pp. 1652-1658, Cape Code, MA, 1996.
- [Albus, 2001] Albus, J. A., *Engineering of Mind: An Introduction to the Science of Intelligent Systems*, pp.195 – 239, John Wiley & Sons, New York, 2001.
- [Allen, 1999] Allen, J.F. Mixed-Initiative Interaction, *IEEE Intelligent Systems*, 14(6), pp.14 – 16, 1999.
- [Amai et al., 2000] Amai, W., J. C. Fahrenholtz, and C. Leger, Hands-free Operation of a Small Mobile Robot, presented at *IEEE Workshop on Vehicle Teleoperation Interfaces*, San Francisco, CA, 2000.
- [Arkin, 1998] Arkin, R., *Behavior-based Robotics*, pp. 491, MIT Press, Cambridge, 1998.
- [Banks et al., 1997] Banks, R., C.D. Wickens, and S. Hah, Commander's Display of Terrain Information: Manipulations of Display Dimensionality and Frame of Reference to Support Battlefield Visualization, Technical Report ARL-97-twelve/ARMY FED LAB 97-2, Savoy, IL, University of Illinois Institute of Aviation, 1997.
- [Bares et al., 1997] Bares, John E. and Wettergreen, David S. Lessons from the Development and Deployment of Dante II, Proceedings of the 1997 *Field and Service Robotics Conference*, Canberra, Australia, December 1997.
- [Barnes et al., 1998] Barnes, M. J., and C. D. Wickens, The Commander's Ability to Visualize Battle spaces: A Multi-View Approach, in Proceedings of 2<sup>nd</sup> *Annual Fedlab Symposium, Advanced Display and Interactive Displays Consortium*, pp. 1-5, College Park, MD, 1998.
- [Bartneck, 2001] Bartneck, C., and, M. Okada, Robofesta – Robotic User Interfaces in Japan, presented at *Symposium on Multimodal Communication with Embodied Agents*, CWI, Amsterdam, The Netherlands, 2001.

- [Bejczy et al., 1990] Bejczy, A., W. Kim, and S. Venema, The phantom robot: Predictive display for teleoperation with time delay, in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp.546-551, Cincinnati, OH, 1990.
- [Berman, 2002] Berman, E.M., *Essential Statistics for Public Managers and Policy Analysts*, 200 pp., CQ Press, Washington, DC, 2002.
- [Bevan et al., 1991] Bevan, N., J. Kirakowski, and J. Maissel, What is Usability?, in *Proceedings of the 4th International Conference on Human Computer Interaction*, Stuttgart, Germany, 1991.
- [Bevan et al., 1994] Bevan, N. and, M. Macleod, Usability measurement in context, *Behaviour and Information Technology*, 13(1-2), pp.132 – 145, 1994.
- [Bevan et al., 1997 ] Bevan, N. and, I. Curson, Methods for Measuring Usability, in *Proceedings of the 6th International Federation for Information Processing Conference on Human-Computer Interaction*, pp. 672 – 673, Sydney, Australia, 1997.
- [Bevan, 1995a] Bevan, N., Measuring usability as quality of use, *Journal of Software Quality Issue*, 4, pp. 115-140, 1995.
- [Bevan, 1995b] Bevan, N., Human-Computer Interaction Standards, in *Proceedings of the 6th International Conference on Human Computer Interaction*, Yokohama, Japan, 1995.
- [Bevan, 1995c] Bevan, N., Usability is quality of use, in *Proceedings of the 6<sup>th</sup> International Conference on Human Computer Interaction*, Yokohama, Japan, 1995.
- [Beveridge et al., 1995] Beveridge, J. R., A. Hanson, and D. Panda, Model-based Fusion of FLIR, Color and LADAR, in *Proceedings of the SPIE-The International Society for Optical Engineering, Sensor Fusion and Networked Robotics VIII*, vol. 2589, P.S. Schenker, and G.T. McKee (Eds.), pp. 2 –11, Philadelphia, PA, 1995.
- [Bluman, 1998] Bluman, A.G., *Elementary Statistics A Step by Step Approach*, 3<sup>rd</sup> Edition, pp. 856, McGraw-Hill, New York, NY, 1998.
- [Boles et al., 2001a] Boles, D. P., and L. P. Adair, The Multiple Ratings Questionnaire (MRQ), in *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, vol. 45, pp. 1790 – 1794, 2001.
- [Boles et al., 2001b] Boles, D. P., and L. P. Adair, The Validity of the Multiple Ratings Questionnaire (MRQ), in *Proceedings of the Human Factors and Ergonomics Society 45<sup>th</sup> Annual Meeting*, vol. 45, pp. 1795-1799, 2001.
- [Borenstein et al., 1996] Borenstein, J., H. R. Everett, and L. Feng, *Navigating Mobile Robots: Systems and Techniques*, pp. 282, A. K. Peters, Wellesley, 1996.
- [Bowman, 1999] Bowman, D. A., Interaction Techniques for common tasks in immersive virtual environments: design, evaluation, and application, Ph.D. thesis, Computer Science, Georgia Institute of Technology, Carnegie Mellon University, Atlanta,

Georgia, 1999.

- [Brown, 1988] Brown, C., *Human-Computer Interface Design Guidelines.*, Ablex Publishing Corporation, Norwood, 1988.
- [Bullinger et al., 1999a] Bullinger, H. and J. Ziegler, (Eds.), *Human-Computer Interaction: Ergonomics and user interfaces*, Proceedings of the 8<sup>th</sup> International Conference on Human-Computer Interaction, vol. 1, Munich, Germany, 1999.
- [Bullinger et al., 1999b] Bullinger, H. and J. Ziegler, (Eds.), *Human-Computer Interaction: Communication, Cooperation, and Application Design*, Proceedings of HCI International, 8<sup>th</sup> International Conference on Human-Computer Interaction, vol. 2, Munich, Germany, 1999.
- [Capocaccia et al., 1988] Capocaccia, G., A., Damasio, C. Regazzoni, G. Vernazza, Data Fusion Approach to obstacle detection and identification, in Proceedings of the SPIE-The International Society for Optical Engineering: Sensor Fusion: Spatial Reasoning and Scene Interpretation, 1003, P.S. Schenker (Ed.), pp. 409-419, Cambridge, Massachusetts, 1988.
- [Carroll, 1997] Carroll, J. M., Human-computer interaction: psychology as a science of design, *International Journal of Human-Computer Studies*, pp. 46, 501-522, 1997.
- [Carroll, 2002] Carroll, J. (Ed.), *Human-Computer Interaction in the New Millennium*, Addison-Wesley, New York, 2002.
- [Cha et al., 1997] Cha, D.-W. and P. Park, Simulator-Based Mental Workload Assessment of the In-Vehicle Navigation System Driver Using Revision of NASA-TLX, *IE Interfaces*, 10(1), pp. 145-154, 1997.
- [Cha et al., 2001] Cha, D.-W. and P. Park, Comparative Study of Subjective Mental Workload Assessment Techniques for the Evaluation of ITS-oriented Human-Machine Interface Systems, *Journal of Korean Society of Transportation*, 19(3), pp.45-58, 2001.
- [Chin et al., 1988] Chin, J. P., V. A. Diehl, and K. L. Norman, Development of a tool measuring user satisfaction of the human-computer interface, in Proceedings of *Computer Human Interaction (CHI) '88: Human Factors in Computing*, E. Soloway, D. Frye, and S. B. Sheppard (Eds.), pp. 213-218, Washington, DC, 1988.
- [Cohen et al., 1982] Cohen, P. R., and E. A. Feigenbaum (Eds.), *The Handbook of Artificial Intelligence.*, 3, pp. 639., William Kaufmann, Inc., Los Altos, 1982.
- [Cohen, 1995] Cohen, P. R., *Empirical Methods for Artificial Intelligence*, pp. 405, The MIT Press, Cambridge, 1995.
- [Davis, 1989] Davis, F. D., Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, *MIS Quarterly*, 13(3), pp. 319 – 340, 1989.

- [Doherty et al., 2001] Doherty, S. M. and C. D. Wickens, Effects of Preview, Prediction, Frame of Reference, and display gain in tunnel in the sky displays, in *Proceedings of the 11<sup>th</sup> International Symposium on Aviation Psychology*, Columbus, OH, 2001.
- [Dorais et al., 1998] Dorais, G., R. P. Bonasso, D. Kortenkamp, B. Pell and D. Schreckenghost, Adjustable Autonomy for Human-Centered Autonomous Systems on Mars, in *Proceedings of the 1<sup>st</sup> International Mars Society Convention*, Boulder, CO, 1998.
- [Dorais et al., 2001] Dorais, G. and D. Kortenkamp, Designing Human-Centered Autonomous Agents, *Pacific Rim International Conference on Artificial Intelligence Workshop Reader, Lecture Notes in Artificial Intelligence 2112*, R. Kowalczyk, S. W. Lake, N. Reed, and G. Williams (Eds.), Springer-Verlag, New York:, 2001.
- [Draper, 2000a] Draper, M. H. and H. A. Ruff, Multi-Sensory Displays and Visualization Techniques Supporting the Control of Unmanned Air Vehicles, presented at *IEEE Workshop on Vehicle Teleoperation Interfaces*, San Francisco, 2000.
- [Drotning et al., 1991] Drotning, W., B. Christensen and S. Thunborg, Graphical model based control of intelligent robot systems, in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 1377 – 1383, Sacramento, CA, 1991.
- [Dumas et al., 1999] Dumas, J.S. and J. C. Redish, *A Practical Guide to Usability Testing*, pp. 416, Intellect Books, Exeter, England, 1999.
- [Eberts, 1994] Eberts, R., *User Interface Design*, pp. 649, Prentice-Hall, Englewood Cliffs, NJ, 1994.
- [Edmonds, 1992] Edmonds, E. (Ed.), *The Separable User Interface: Computers and People Series*, pp. 371, New York: Academic Press, New York, 1992.
- [Endsley, 1989] Endsley, M. R., A Methodology for the objective measurement of Pilot Situation Awareness, in *Advisory Group for Aerospace Research and Development (AGARD) Conference Proceedings*, vol. 478, pp. 1 – 9, Copenhagen, Denmark, 1989.
- [Endsley, 1995] Endsley, M. R.. Toward a Theory of Situation Awareness in Dynamic Systems, *Human Factors*, 37(1), pp. 32 – 64, 1995.
- [Everett, 1995] Everett, H. R., *Sensors for Mobile Robots: Theory and Application*, pp. 544, A. K. Peters, Wellesley, 1995.
- [Fijalkiewicz, 1998] Fijalkiewicz, P., and G. Dejong, Cheshire: An Intelligent Adaptive User Interface, in *Proceedings of 2<sup>nd</sup> Annual Fedlab Symposium, Advanced Display and Interactive Displays Consortium*, pp. 15 –19, College Park, MD, 1998.
- [Flach, 1995] Flach, J. M., Situation Awareness: Proceed with Caution, *Human Factors*, 37(1), pp. 149-157, 1995.
- [Fleming et al., 2000] Fleming, M.C., and J.G. Nellis, *Principles of applied statistics: an integrated*

- approach using Minitab and Excel*, 2<sup>nd</sup> Edition, pp. 480, Thomson Learning, London, England, 2000.
- [Fong et al., 2000] Fong, T., F. Conti, S. Grange, and C. and Baur, Novel interfaces for remote driving: gesture, haptic and PDA, in *Proceedings of the SPIE - The International Society of Optical Engineering, Telemanipulator and Telepresence VII*, vol. 4195, Boston, MA, 2000.
- [Fong et al., 2001a] Fong, T and C. Thorpe, Vehicle Teleoperation Interfaces, *Autonomous Robots*, 11(1), pp. 9 – 18, 2001.
- [Fong et al., 2001b] Fong, T., C. Thorpe, and C. Baur, Advanced Interfaces for Vehicle Teleoperation: Collaborative Control, Sensor Fusion Displays, and Remote Driving Tools, *Autonomous Robots*, 11(1), pp. 77 – 85, 2001.
- [Fong et al., 2001c] Fong T., C. Thorpe, and C. Baur, Active Interfaces for Vehicle Teleoperation, in *Proceedings of the SPIE – The International Society for Optical Engineering, Robotics and Machine Perception Working Group Newsletter*, 10(1), 2001.
- [Fong, 2001] Fong, T., Collaborative control: A Robot-Centric model for Vehicle Teleoperation, Ph.D. Thesis, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA., 2001.
- [Fredslund, 2002] Fredslund, J., and M. J. Mataric, Hewie, Dewie, Louie, and GUI-Commanding Robot Formation, in *Proceedings of the 2002 IEEE International Conference on Robotics and Automation*, pp. 175 – 180, Washington DC, 2002.
- [French, 1995] French, R. L., From Chinese Chariots to Smart Cars: 2000 Years of Vehicular Navigation, *Journal of the Institute of Navigation*, 42(1), pp. 235-258, 1995.
- [Frokjaer et al., 2000] Frokjaer, E., M. Hertzum, and K. Hornbaek, Measuring Usability: Are Effectiveness, Efficiency and Satisfaction Really Correlated?, in the *Proceedings of the Computer-Human Interaction 2000*, pp. 345-352, T. Turner, G. Szwillus, M. Czerwinski, and F. Paterno (Eds.), ACM, The Hague, Netherlands, 2000.
- [Fu, et al., 1987] Fu, K.S., R. C. Gonzalez, and C. S. G. Lee, *Robotics: Control, Sensing, Vision and Intelligence*, pp. 580, McGraw-Hill, St. Louis, 1987.
- [Gediga et al., 2001] Gediga, G. , K. Hamborg, and I. Duntsch, Evaluation of Software Systems, *Encyclopedia of Computer Science and Technology*, 45, A. Kent and J. G. Williams (Eds.), 2001.
- [Gibson, 1950] Gibson, J., *Perception of the Visual World*, Houghton Mifflin Company, Boston, 1950.
- [Gilbreath et al., 2001] Gilbreath, G. A., D. A. Ciccimaro, and H. R. Everett, An Advanced Telereflexive Tactical Response Robot, *Autonomous Robots*, 11(1), 2001.

- [Gourley et al., 1997] Gourley, C. S., and M. A. Abidi, Virtual Reality Hardware for Use in Interactive 3-D Data Fusion and Visualization, in *Proceedings of the SPIE – The International Society for Optical Engineering, Sensor Fusion and Decentralized Control in Autonomous Robotic Systems*, vol. 3209, P. S. Schenker and G. T. McKee (Eds.), pp. 281-291, Pittsburgh, PA, 1997.
- [Graefe, 1998] Graefe, V., Perception and Situation Assessment for Behavior-Based Robot Control, *Intelligent Autonomous Systems*, Y. Kakazu, N. Wada (Eds.), and T. Sato, pp. 376-383, 1998.
- [Grissom et al., 1993] Grissom, S., Perlman, G., StEP(3-D): A portable discount usability evaluation plan for 3-D interaction. Ohio State University, Department of Computer Science and Information Science, Technical Report OSU-CISRC-2/93-TR7, 1993.
- [Hainsworth, 2001] Hainsworth, D., Teleoperation user interfaces for mining robotics, *Autonomous Robots*. 11(1), pp. 19 – 28, 2001.
- [Halme et al., 2000] Halme, A., and J. Suomela, Tele-existence Techniques of Heavy Work Vehicles, in *Proceedings of IEEE International Conference on Robotics and Automation*, San Francisco, CA, pp. 29-37, 2000.
- [Hancock et al., 1988] Hancock, P., and N. Meshkati, *Human Mental Workload*., Elsevier Science Publishing Company, New York, 1988.
- [Hart et al., 1988] Hart, S.G. and Staveland, L.E., Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, in *Human Mental Workload*, P.A. Hancock and N. Meshkati (Eds.), pp. 139-183, Elsevier Science Publishing Company, New York, 1988.
- [Hebert, 2000] Hebert, M., Active and Passive Range Sensing for Robotics, in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 1, pp. 102 – 110, San Francisco, CA, 2000.
- [Henderson et al., 1988] Henderson, T. C., and O. D. Faugeras, High-Level Multisensor Integration, in *Proceedings of the SPIE-The International Society for Optical Engineering: Sensor Fusion: Spatial Reasoning and Scene Interpretation*, 1003, P.S. Schenker (Ed.), pp. 307-314, Cambridge, MA, 1988.
- [Hill et al., 1992] Hill, S. G., H. P. Iavecchia, J. C. Byers, A. C. Bittner, A. L. Zaklad, and R. E. Christ, Comparison of Four Subjective Workload Rating Scales, *Human Factors*, 34(4), pp. 429 – 439, 1992.
- [Hix et al., 1993] Hix, D. and H. Rex, *Developing User Interfaces: Ensuring Usability through Product and Process*, pp. 416, John Wiley and Sons, New York, 1993.
- [Holtzblatt et al., 1996] Holtzblatt, K. and H. Beyer, Contextual Design: Principles and Practice, *Field Methods for Software and Systems Design*., D. Wixon and J. Ramey

- (Eds.), pp. 301-333, John Wiley and Sons, New York, 1996.
- [Horiguchi, 2000] Horiguchi, Y, T. Sawaragi, and G. Akahi, Naturalistic Human-Robot Collaboration Based Upon Mixed-Initiative Interactions in Teleoperating Environment, in *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, vol. 2, pp. 876-881, Nashville, TN, 2000.
- [Horvitz, 1999] Horvitz, E., Principles of Mixed-Initiative User Interfaces, in *Proceedings of Computer Human Interaction '99, ACM Conference on Human Factors in Computing Systems*, pp. 159 – 166, Pittsburgh, PA, 1999.
- [Hu et al., 1997] Hu, H., D. Gu, and M. Brady, Outdoor Navigation of a Mobile Robot with Multiple Sensors, in *Proceedings of the SPIE – The International Society for Optical Engineering, International Symposium on Intelligent Systems & Advanced Manufacturing, Mobile Robots XII*, pp. 13-24, 1997.
- [Hughes, ] Hughes, S., and Lewis, M., Attentive Camera Navigation in Virtual Environments, in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 967 – 970, Nashville, TN, 2000.
- [Jacob, 1994] Jacob, R .J. K., New Human-Computer Interaction Techniques, *Human-Machine Communication for Education Systems Design*, M. D. Brouwer-Janse, and T. L. Harrington (Eds.), pp. 131 – 138, Springer-Verlag, Berlin, 1994.
- [Jacob, 1996] Jacob, R. J. K., Input Devices and Techniques, *The Computer Science and Engineering Handbook*, A. B. Tucker (Ed.), pp. 1494-1511, CRC Press, Boca Raton, LA, 1996.
- [Jacob, 2000] Jacob, R. J. K., User Interfaces, *Encyclopedia of Computer Science*, 4th edition, A. Ralston, E. D. Reilly, and E. Hemmendinger (Eds.), Grove Dictionaries, Inc., New York, 2000.
- [Johnson, 2000] Johnson, R., *Miller and Freund's Probability and Statistics for Engineers*, pp. 622, Prentice Hall, Upper Saddle River, NJ, 2000.
- [Johnson, 2002] Johnson, C. A., A. B. Koku, K. Kawamura, and R.A. Peters II, Enhancing a human-robot interface using Sensory EgoSphere, in *Proceedings of the 2002 IEEE International Conference on Robotics and Automation*, 2002, Washington DC, pp. 4132-4137, 2002.
- [Jones et al., 1999] Jones, J. L., G. Seiger, and A. M. Lynn, *Mobile Robots: Inspiration to Implementation*, pp. 486, A.K. Peters Ltd, Natick, MA, 1999.
- [Jones, 1989] Jones, M., *Human-Computer Interaction: A design guide.*, Educational Technology Publications, Englewood Cliffs, NJ, 1989.
- [Kawamura et al., 2000] Kawamura, K., R. A. Peters II, D. M. Wilkes, A. Alford, and T. E. Rogers, ISAC: Foundations in Human-Humanoid Interaction, *IEEE Intelligent*

*Systems*, 15(4), pp. 38-45, 2000.

- [Kawamura et al., 2001a] Kawamura, K., R.A. Peters II, C. A. Johnson., P. Nilas, and S. Thongchai, Supervisory Control of Mobile Robot using Sensory EgoSphere in Proceedings of 2001 *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pp. 523 – 529, Banff, Canada, 2001.
- [Kawamura et al., 2001b] Kawamura, K., R. A. Peters II, A. B. Koku, and A. Sekmen, Landmark EgoSphere-based topological navigation of mobile robots, in Proceedings of the *SPIE – The International Society for Optical Engineering, Intelligent Systems and Advance Manufacturing*, vol. 4573, pp., Newton, MA, 2001.
- [Kawamura et al., 2001c] Kawamura, K., R. A. Peters II, D. M. Wilkes, A.. B. Koku, and A. Sekmen, Toward Perception-Based Navigation using EgoSphere, in Proceedings of the *SPIE – The International Society for Optical Engineering, Intelligent Systems and Advanced Manufacturing*, vol. 4573, pp.137-147, Newton, MA, 2001.
- [Kawamura et al., 2002a] Kawamura, K., D. M. Wilkes, A. B. Koku, and T. Keskinpala, Perception-Based Navigation for Mobile Robots, presented at *Multi-robot systems Workshop*, Naval Research Laboratory, Washington, DC, 2002.
- [Kawamura et al., 2002b] Kawamura, K., A.B. Koku, D.M. Wilkes, R.A. Peters II, and A. Sekmen, Toward Egocentric Navigation, *International Journal of Robotics and Automation*, 17(4), pp.135-145, November 2002.
- [Klingspor, 1997] Klingspor, V., J. Demiris, and M. Kaiser, Human-Robot Communication and Machine Learning, *Applied Artificial Intelligence Journal*, 11, pp. 719-746, 1997.
- [Koku et al., 1999] Koku, A. B., and R. A. Peters II, A Data Structure for the Organization by a Robot of Sensory Information, in Proceedings of the 2<sup>nd</sup> *International Conference on Recent Advances in Mechatronics*, Istanbul, Turkey, 1999.
- [Koren, 1985] Koren, Y., *Robotics for Engineers*, McGraw-Hill, St. Louis, MO, 1985.
- [Kortenkamp, 1997] Kortenkamp, D., R. P. Bonasso, D. Ryan, and D. Schreckenghost, Traded Control with Autonomous Robots as Mixed Initiative Interaction, *AAAI Spring Symposium on Mixed Initiative Interaction*, Technical Report SS-97-04, pp. 89-94, March 1997.
- [Kortenkamp, 2000] Kortenkamp, D., D. Keirn-Schreckenghost, and R. P. Bonasso, Adjustable Control Autonomy for Manned Space Flight, in Proceedings of *IEEE Aerospace Conference*, Big Sky, Montana, 2000.
- [Lane et al., 2001] Lane, J. C., C. R. Carignan, and D. L. Akin, Advanced Operator Interface Design for Complex Space Telerobots, *Autonomous Robots*, 11(1), pp. 49–58, 2001.
- [Langley, 1992] Langley, P. User Modeling in Adaptive Interfaces, in Proceedings of the

- 7th International Conference on User Modeling, pp. 357-370, Springer, Banff, Alberta, 1992.
- [Lewis, 1995] Lewis, J. R. IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use, *International Journal of Human-Computer Interaction*, 7(1), pp. 57 – 78, 1995.
- [Liebowitz et al., 1987] Liebowitz, S., and D. Casasent., Multisensor processing: object detection and identification, in Proceedings of the *SPIE – The International Society for Optical Engineering, Mobile Robots II*, vol. 852, pp. 54- 71, 1987.
- [Lin et al., 1997] Lin, H.X., Y.-Y. Choong, and G. Salvendy, A Proposed Index of Usability: A Method for Comparing the Relative Usability of Different Software Systems, *Behavior & Information Technology*, 16(4/5), pp. 267 – 278, 1997.
- [Liu et al., 2000a] Liu, J. S. Pastoor, K. Seifert, and J. Hurtienne, Three-dimensional PC: toward novel forms of human-computer interaction, in Proceedings of the *SPIE – The International Society for Optical Engineering, Photonics East Critical Review Conference on Three-dimensional Video and Display*, Boston, MA, pp. 250-281, 2000.
- [Liu et al., 2001b] Liu, Q., T. Huang, Y. Wu, and S. Levinson, Spoken language acquisition via human-robot interaction, in Proceedings of *IEEE International Conference on Multimedia and EXPO*, Tokyo, Japan, 2001.
- [Macleod et al., 1997] Macleod, M., R. Bowden, N. Bevan, and I. Curson, The MUSiC Performance Measurement Method, *Behaviour and Information Technology*, 16, pp.279 – 293, 1997.
- [Macleod, 1993] Macleod, M. and R. Rengger, The Development of DRUM: A Software Tool for Video-assisted Usability Evaluation, *People and Computers VIII, Proceedings of Human Computer Interaction '93 Conference*, Loughborough UK, pp. 293-309, 1993.
- [Macleod, 1994] Macleod, M., Usability: Practical Methods for Testing and Improvement, in Proceedings of the *Norwegian Computer Society Software 94 Conference*, Sandvika, Norway, February 1994.
- [Malcolm, 1988] Malcolm, D. R., *Robotics: An Introduction.*, pp. 368, PWS-Kent Publishing Co., Boston, MA, 1998
- [Mansour et al., 1992] Mansour, R., and W. Karwowski, *Human-Robot Interaction.*, Taylor & Francis, Washington, DC, 1992.
- [Marsh et al., 1999] Marsh, T. and P. Wright, Cooperative Evaluation of a Desktop Virtual Reality System, presented at *1999 Workshop on User Centered Design and Implementation of Virtual Environment*, King's Manor, University of York, 1999.
- [Matsui et al., 1990] Matsui, T. and M. Tsukamoto, Integrated robot teleoperation method using multi-media display, in Proceedings of the *5<sup>th</sup> International Symposium on*

- Robotics Research*, pp. 145-152, 1990.
- [Meier et al., 1999] Meier, R., T. Fong, C. Thorpe, and C. Baur, A Sensor Fusion Based User Interface for Vehicle Teleoperation, in Proceedings of the *IEEE Field and Service Robotics*, Pittsburgh, PA, 1999.
- [Mendenhall, 1987] Mendenhall W, *Introduction to Probability and Statistics*, 7<sup>th</sup> edition, Duxbury Press, Boston, MA, 1987.
- [Merritt, 1988] Merritt, J. O., Virtual Window viewing geometry, in Proceedings of *SPIE-The International Society for Optical Engineering: Sensor Fusion: Spatial Reasoning and Scene Interpretation*, 1003, P.S. Schenker (Ed.), pp. 386-389, Cambridge, MA, 1988.
- [Milgram et al., 1993] Milgram, P., S. Zhai, S. and D. Drascic, Applications of Augmented Reality for Human-Robot Communication, in Proceedings of the *1993 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1467-1472, Yokohama, Japan, 1993.
- [Miltonberger et al., 1988] Miltonberger, T., D. Morgan, and G. Orr, Multisensor object recognition for 3-D models, in Proceedings of the *SPIE-The International Society for Optical Engineering: Sensor Fusion: Spatial Reasoning and Scene Interpretation*, 1003, P.S. Schenker (Ed.) , pp. 161-169, Cambridge, MA, November 1988.
- [Moore, 1995] Moore, D.S., *The Basic Practice of Statistics*, pp. 688, W.H. Freeman And Company, New York, NY, 1995.
- [Murphy et al., 1996] Murphy, R., K. Gome, and D. Hershberger, Ultrasonic Data Fusion as a Function of robot velocity, in Proceedings of the *SPIE-The International Society for Optical Engineering: Sensor Fusion and Distributed Robotic Agents*, 2905, P.S. Schenker, and G.T. McKee (Eds.), pp. 114-126, Boston, MA, November 1996.
- [Murphy, 1996] Murphy, R., and E. Rogers, Cooperative Assistance for Remote Robot Supervision, *Presence, special issue on Starkfest*, 5(2), pp. 224-240, 1996.
- [Murphy, 2000] Murphy, R., *Introduction to AI Robotics*, pp. 400, The MIT Press, Cambridge, MA, 2000.
- [Murphy et al., 2001] Murphy, R., J. Casper, M. Micire, and J. Hyams, Mixed-Initiative Control of Multiple Heterogeneous Robots for Urban Search and Rescue, submitted to *IEEE Transactions on Robotics and Automation*, 2001.
- [Nash, 2000] Nash, E. B., The Effect of Communication Style on Task Performance and Mental Workload Using Wearable Computers, M.S. thesis, Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 2000.
- [Nguyen et al., 2001] Nguyen, L. M. Bualat, L. J. Edwards, L. Flueckiger, C. Neveu, K. Schwehr, K., M. D. Wagner, and E. Zbinden, Virtual Reality Interfaces For Visualization and Control of Remote Vehicles, *Autonomous Robots*, 11(1), pp. 59 –

68, 2001.

- [Nielsen, 1992] Nielsen, J., Finding Usability Problems through Heuristic Evaluation, in Proceedings of the *Computer Human Interaction 1992*, P. Abuersfeld, J. Bennett, and G. Lynch (Eds.), pp. 373-380, Monterey, CA, May 1992.
- [Nielsen, 1993] Nielsen, J., *Usability Engineering*, pp. 362, AP Professional, Boston, MA, 1993.
- [Nilas, 2003] Nilas, P., Multi-Agent Based Adaptive Human-Robot Interface, Ph.D. thesis, Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, May 2003.
- [Nishiuchi et al., ] Nishiuchi, S., K. Kurihara, S. Sakai, and H. Takada, Using human line-of-sight to control equipment, in Proceedings of *I MECH E Part B Journal of Engineering Manufacture*, 215(4), pp. 577 – 582, 2001.
- [Norman et al., 1986] Norman, D., and S. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1986.
- [Norman, 1993] Norman, D., *Things that make us smart: defending human attributes in the age of the machine*, pp. 253, Perseus Books, Cambridge, MA, 1993.
- [Norusis, 2002] Norusis, M.J., *SPSS 11.0: Guide to Data Analysis*, pp. 637, Prentice Hall, Upper Saddle River, NJ, 2002.
- [Pack, 1998] Pack, R. T., IMA: The Intelligent Machine Architecture, Ph.D. Thesis, Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, 1998.
- [Park et al., 2001] Park, H., B. Kim, and K. Lim, Measuring the Machine Intelligence Quotient (MIQ) of Human-Machine Cooperative Systems, *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 31(2), pp. 89- 96, 2001.
- [Paulos et al., 2001] Paulos, E. and J. Canny, Social Tele-Embodiment: Understanding Presence, *Autonomous Robots*, 11(1), pp. 87-95, 2001.
- [Perzanowski, 1999] Perzanowski, D., A. C. Schultz, W. Adams, and E. Marsh, Goal Tracking in a Natural Language Interface: Towards Achieving Adjustable Autonomy, in Proceedings of the *IEEE International Symposium, on Computational Intelligence in Robotics and Automation*., pp. 208-213, Monterey, CA, 1999.
- [Pesante-Santana, 1997] Pesante-Santana, J. A., The Effects of Multitasking on Quality Inspection in Advanced Manufacturing Systems, Ph.D. Thesis, Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 1997.
- [Peters et al., 2001] Peters II, R.A., K. A. Hambuchen, K., Kawamura, and D. M. Wilkes, The Sensory EgoSphere as a Short-Term Memory for Humanoids, in Proceedings

- of *IEEE-Robotics and Automation Society International Conference on Humanoid Robots*, pp. 451-459, Tokyo, Japan, 2001.
- [Peters, 2001] Peters, R.A., Sensory Processing Module Linkage to SES, Working papers, Vanderbilt University, 2001.
- [Piguet, 1995] Piguet, L., T. W. Fong, B. Hine, P. Hontalas, and E. Nygren, VEV: A Virtual Reality Tool For Robotic Planetary Exploration, in *Proceedings of Virtual Reality World*, pp. 273, Stuttgart, Germany, 1995.
- [Preece et al., 1994] Preece, J., Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey, *Human-Computer Interaction*, pp. 492, Addison-Wesley, New York, 1994.
- [Prothero, 1994] Prothero, J., Survey of Interface Goodness Measures, March 16, 1994, University of Washington, Human Interface Technology Lab Technical Report R-94-1. <http://www.hitl.washington.edu/publications/r-94-1/r-94-1.pdf>
- [Pugh, 1986] Pugh, A., *Robot Sensors Volume 2: Tactile and Non-Vision*, pp. 450, Springer-Verlag, New York, 1986.
- [Rahmini et al., 1992] Rahmini, M., and W. Karwowski, (Eds.), *Human-Robot Interaction*, Taylor And Francis, Washington, DC, 1992.
- [Ravden et al., 1989] Ravden, S. and G. Johnson, *Evaluating usability of Human-Computer Interfaces: A Practical Method User Interface Design*, John Wiley & Sons, New York, 1989.
- [Riley, 1989] Riley, V., A general model of mixed-initiative human-machine systems, in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, pp. 124-128, Denver, CO, 1989.
- [Rossman et al., 1997] Rossman, A.J., and J.B. Von Oehsen, *Workshop Statistics: Discovery with Data and the Graphing Calculator*, 1st Edition, pp. 624, Springer -Verlag, New York, NY, 1997.
- [Rosson et al., 2002] Rosson, M.B. and J. M. Carroll, *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*, pp. 448, Morgan Kaufmann Publishers, New York, 2002.
- [Rouse et al., 1993] Rouse, W.B., S. L. Edwards, and J. M. Hammer, Modeling the Dynamics of Mental Workload and Human Performance in Complex Systems, *IEEE Transactions on Systems, Man and Cybernetics*, 23(6), pp. 1662-1671, 1993.
- [Salas et al., 1995] Salas, E., C. Prince, D. P. Baker, and L. Shrestha, Situation Awareness in Team Performance: Implications for Measurement and Training, *Human Factors and Ergonomics Society*, 37(1), pp. 123-136, 1995.
- [Sayers et al., 1994] Sayers, C., and R. P. Paul, An Operator Interface for Teleprogramming Employing Synthetic Fixtures, *Presence*, 3(4), pp. 309- 320, 1994.

- [Sayers, 1999] Sayers, C., *Remote Control Robotics*, pp. 240, Springer-Verlag, New York, 1999.
- [Scholtz, 2001] Scholtz, J., Adaptation of Traditional Usability Testing Methods for Remote Testing, in *Proceedings of the 34<sup>th</sup> Annual Hawaii International Conference on System Science*, pp.5030, 2001.
- [Scholtz, 2002] Scholtz, J., Human Robot Interactions: Creating Synergistic Cyber Forces, *AAAI Fall Symposium on Human-Robot Interaction, Technical Report FS-02-03*, pp. 81 – 91, 2002.
- [Schuler et al., 1993] Schuler, D. and A. Namioka (Eds.), *Participatory Design*, pp. 312 pp., Lawrence Erlbaum Associates, Hillsdale, NJ, 1993.
- [Sekmen, 2000] Sekmen, A. S., Human-robot interaction methodology, Ph.D. Thesis, Electrical and Computer Engineering, Vanderbilt University, 2000.
- [Selcon et al., 1989] Selcon, S. J. and R. M. Taylor, Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design, in *Proceedings of the Advisory Group for Aerospace Research and Development Conference Proceedings*, vol. 478, pp.5/1 – 5/8, Copenhagen, Denmark, 1989.
- [Shepard et al., 1971] Shepard, R. N. and J. Metzler, Mental Rotation of Three-Dimensional Objects, *Science*, New Series, 171(3972), pp. 701-703, 1971.
- [Sheridan, 1992] Sheridan, T., *Telerobotics, Automation and Human Supervisory Control*, pp. 432, The MIT Press, Cambridge, MA, 1992.
- [Shneiderman, 1987] Shneiderman, B., *Designing the User Interface*, pp. 640, Addison-Wesley, Reading, PA, 1987.
- [Sincich et al., 1999] Sincich, T.L., D.M. Levine, and D. Stephan, *Practical Statistics by Example using Microsoft Excel*, pp. 789, Prentice Hall, Upper Saddle River, NJ, 1999.
- [Skubic et al., 2002] Skubic, M., D. Perzanowski, A. Schultz, and W. Adams, Using Spatial Language in a Human-Robot Dialog, in *Proceedings of the 2002 IEEE International Conference on Robotics and Automation*, vol. 4, pp. 4143 – 4148, Washington, DC, 2002.
- [Smith et al., 1995] Smith, K. and P. A. Hancock, Situation Awareness is Adaptive, Externally Directed Consciousness, *Human Factors and Ergonomics Society*, 37(1), pp. 1378 – 148, 1995.
- [Stansfield, 1988] Stansfield, S. A., Integrating Multiple Views into a Single Representation of a Range Imaged Object, in *Proceedings of the SPIE-The International Society for Optical Engineering: Sensor Fusion: Spatial Reasoning and Scene Interpretation*, vol. 1003, P.S. Schenker (Ed.), pp. 52 – 62, Cambridge, MA, 1988.
- [Sutcliffe, 1989] Sutcliffe, A., *Human-Computer Interface Design*, Springer-Verlag, New York,

1989.

- [Terrien et al., 2000] Terrien, G., T. Fong, C. Thorpe, and C. Baur, Remote driving with a multisensor user interface, in *Proceedings of the 30<sup>th</sup> Annual Society of Automotive Engineers International Conference on Environmental Systems*, Toulouse, France, 2000.
- [Thongchai, 2001] Thongchai, S., *Intelligent Control and Learning Techniques for Mobile Robots*, Ph.D. thesis, Electrical and Computer Engineering, Vanderbilt University, 2001.
- [Tsang et al., 1996] Tsang, P.S., and V. L. Velazquez, Diagnosticity and multidimensional subjective workload ratings, *Ergonomics*, 39, pp. 358 –381, 1996.
- [Van Orden, 2001] Van Orden, K. F., *Monitoring Moment-to-Moment Operator Workload Using Task Load and System-State Information*, Technical Report 1864, SPAWAR Systems Center, San Diego, CA, 2001.
- [Vandenberg et al., 1979] Vandenberg, S. G., and A. R. Kuse, Mental Rotation: A group test of three-dimensional spatial visualization, *Perceptual and Motor Skills*, 47, pp. 599 – 604, 1979.
- [Wadsworth, 1998] Wadsworth, H.M., *Handbook of Statistical Methods for Engineers and Scientists*, 2<sup>nd</sup> Edition, pp. 768, McGraw-Hill, New York, NY, 1998.
- [Wettergreen, 1997] Wettergreen, D., M. Bualat, D. Christian, D. Schwehr, H. Thomas, D. Tucker, and E. Zbinden, E., *Operating Nomad During Atacama Desert Trek*, presented at *Field and Service Robotics Conference*, Canberra, Australia, 1997.
- [Wickens et al., 1997] Wickens, C. D., O. Olmos, A. Chudy, and C. Davenport, *Aviation Display Support for Situation Awareness*, University of Illinois Institute of Aviation Technical Report (ARL-97-10/LOGICON-97-2), Savoy, IL, Aviation Research Lab, 1997.
- [Wickens, 1984] Wickens, C. D., *Engineering Psychology and Human Performance*, pp. 573, Charles E. Merrill Publishing Company, Columbus, OH, 1984.
- [Wierwille et al., 1993] Wierwille, W. W. and F. T. Eggemeier, Recommendations for Mental Workload Measurement in a Test and Evaluation Environment, *Human Factors*, 35(2), pp. 263-281, 1993.
- [Wild et al., 2000] Wild, P. J. and R. D. Macredie, Usability Evaluation and Interactive Systems Maintenance, in *Proceedings of 2000 Annual Conference for the Computer-Human Interaction Special Interest Group of the Ergonomics Society of Australia: Interfacing Reality in the New Millennium*, C. Paris, S. Howard, and N. Ozkan (Eds.), Sydney, Australia, 2000.
- [Williams et al., 1993] Williams, J. A. and C. M. Mitchell, Effects of Integrated Flight Path and Terrain Displays on Controlled Flight into Terrain, in *Proceedings of the 1993 International Conference of IEEE Systems, Man, and Cybernetics*, pp. 709-714, Le

Touquet, France, 1993.

- [Yang et al., 1995] Yang, J. and Y. Wu, Detection for Mobile Robot Navigation Based on Multisensor Fusion, in Proceedings of the *SPIE - The International Society of Optical Engineering, Mobile Robots X*, vol. 2591, pp. 182-192, 1995.
- [Yong, 1999] Yong, L. T. and T. E. Kong, The Study of Cooperative Evaluation Approach on Internet Search, in Proceedings of *University of Southern Mississippi Computer Science Student Symposium on Computer Science and IT*, Penang, Malaysia, 1999.
- [Zhai et al., 1992] Zhai, S. and P. Milgram, Human Robot Synergism and Virtual Telerobotic Control, in Proceedings of the *25<sup>th</sup> Annual Conference of the Human Factors Association of Canada*, Hamilton, Ontario, 1992.
- [Zhai, 1991] Zhai, S. and P. Milgram, A telerobotic virtual control system, in Proceedings of the *SPIE - The International Society of Optical Engineering, Cooperative Intelligent Robotics in Space II*, vol. 1612, pp. 311-320, Boston, MA, 1991.