REMOTE EVALUATION OF THE COHERENCE OF INDIRECT MANIPULATION INTERFACE SYSTEMS FOR AGENT-MEDIATED LEGACY DATA

By

Joseph Hughes Schafer

B.S. May 1986, The United States Military Academy, West Point, New York

M.S. May 1997, The George Washington University, Washington, District of Columbia

A Dissertation submitted to

The Faculty of

The School of Engineering and Applied Science

of The George Washington University in partial satisfaction

of the requirements for the degree of Doctor of Science

May 21, 2000

Dissertation directed by

Rachelle Silverman Heller

Professor of Engineering and Applied Science

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY) 21-05-2000	**2. REPORT TYPE** Dissertation, Doctor of Science	**3. DATES COVERED (FROM - TO)** xx-xx-2000 to xx-xx-2000
4. TITLE AND SUBTITLE Remote Evaluation of the Coherence of Indirect Manipulation Interface Systems For Agent-Mediated Legacy Data Unclassified		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Schafer, Joseph H. ;		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME AND ADDRESS Computer Science Department George Washington University 725 23rd Street NW Washington , DC 20052		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS US Army Student Detachment Fort Jackson , SC 29207		**10. SPONSOR/MONITOR'S ACRONYM(S)**
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT A PUBLIC RELEASE US Army Student Detachment Fort Jackson , SC 29207		

13. SUPPLEMENTARY NOTES

A Dissertation submitted to The Faculty of The School of Engineering and Applied Science of The George Washington University in partial satisfaction of the requirements for the degree of Doctor of Science

14. ABSTRACT

Many information systems depend heavily on distributed legacy data sources. These data sources introduce a number of significant problems, especially when the sources must be combined and displayed to remote users. Many researchers have investigated various interface systems, however empirical studies have not been published that examine remote interfaces to distributed heterogeneous data. The purpose of this research is to determine the efficacy of a system that provides a more coherent representation of this distributed data in comparison to a more traditional system for users performing representative tasks. This dissertation presents the results of remote usability experiments in a specific, well-defined context. These web-based experiments empirically determine whether coherence is enhanced through application of the proposed methodology by presenting each interface system and a sequence of representative tasks. The remote evaluation system measures coherence based upon the subject's time to complete each task, the correctness of their answer, and their subjective confidence in that answer. When all tasks have been completed, the respondents complete a usability survey to express their satisfaction with the interface system. The specific research undertaking is to determine whether a system based upon this proposed methodology, the Visual Interface To Agent Mediated Information Networks (VITAMIN) system, is superior to a system based upon a traditional approach, the Java Indirect Manipulation Interface (JIMI) system. VITAMIN was developed to add coherence to the legacy data and JIMI was developed as a control treatment to represent a traditional legacy approach. Both VITAMIN and JIMI were implemented as indirect manipulation interface systems, or non-anthropomorphic interface agents. The interface systems provide intermediate query predicate actions to the Heterogeneous Reasoning and Mediator System, HERMES. HERMES is an agent mediator system in the IMPACT agent architecture that provides access to legacy data sources for this research. The remote usability experiment was enabled with the Java Usability Interface Comparison and Evaluation (JUICE) system, which was also developed for this research. JUICE also uses HERMES to chronicle the experimental results. The experiments used a within-subject randomly counter-balanced design involving 56 subjects. An additional 63 subjects were tested in two phases of pilot experiments. The JUICE system randomly varied the order of the treatments to counterbalance the experiment. The representative tasks were divided into three types. This division was based upon the levels of comprehension, application, and analysis required to answer the task according to Bloom's taxonomy. Statistically significant results were obtained for all four indicators of coherence with VITAMIN providing an improvement over JIMI for each indicator: (1) the number of correctly answered tasks; (2) the user confidence in correctly answered tasks; (3) the time to answer tasks correctly; and (4) the user satisfaction with the interface.

15. SUBJECT TERMS

Remote Evaluation; Mediator; Software Agents; Coherence; Indirect Manipulation; Heterogeneous Legacy Data; Visual Interface; Usability Metrics; IMPACT; HERMES; JIMI; VITAMIN; JUICE; Army War Reserve Prepositioned Stocks

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
			Public Release	295	Fenster, Lynn lfenster@dtic.mil
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER International Area Code Area Code Telephone Number 703 767-9007 DSN 427-9007

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
21-05-2000

2. REPORT TYPE
Dissertation, Doctor of Science

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE Remote Evaluation of the Coherence of Indirect Manipulation Interface Systems For Agent-Mediated Legacy Data

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Joseph H. Schafer, MAJ USA

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Computer Science Department
George Washington University
725 23rd Street NW
Washington, DC 20052

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
US Army Student Detachment
Fort Jackson, SC 29207

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A: APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

13. SUPPLEMENTARY NOTES
A Dissertation submitted to The Faculty of The School of Engineering and Applied Science of The George Washington University in partial satisfaction of the requirements for the degree of Doctor of Science

14. ABSTRACT. **Many information systems depend heavily on distributed legacy data sources. These data sources introduce a number of significant problems, especially when the sources must be combined and displayed to remote users. Many researchers have investigated various interface systems, however empirical studies have not been published that examine remote interfaces to distributed heterogeneous data. The purpose of this research is to determine the efficacy of a system that provides a more coherent representation of this distributed data in comparison to a more traditional system for users performing representative tasks. This dissertation presents the results of remote usability experiments in a specific, well-defined context. These web-based experiments empirically determine whether coherence is enhanced through application of the proposed methodology by presenting each interface system and a sequence of representative tasks. The remote evaluation system measures coherence based upon the subject's time to complete each task, the correctness of their answer, and their subjective confidence in that answer. When all tasks have been completed, the respondents complete a usability survey to express their satisfaction with the interface system.**

The specific research undertaking is to determine whether a system based upon this proposed methodology, the Visual Interface To Agent Mediated Information Networks (VITAMIN) system, is superior to a system based upon a traditional approach, the Java Indirect Manipulation Interface (JIMI) system. VITAMIN was developed to add coherence to the legacy data and JIMI was developed as a control treatment to represent a traditional legacy approach. Both VITAMIN and JIMI were implemented as indirect manipulation interface systems, or non-anthropomorphic interface agents.

The interface systems provide intermediate query predicate actions to the Heterogeneous Reasoning and Mediator System, HERMES. HERMES is an agent mediator system in the IMPACT agent architecture that provides access to legacy data sources for this research. The remote usability experiment was enabled with the Java Usability Interface Comparison and Evaluation (JUICE) system, which was also developed for this research. JUICE also uses HERMES to chronicle the experimental results.

The experiments used a within-subject randomly counter-balanced design involving 56 subjects. An additional 63 subjects were tested in two phases of pilot experiments. The JUICE system randomly varied the order of the treatments to counterbalance the experiment. The representative tasks were divided into three types. This division was based upon the levels of comprehension, application, and analysis required to answer the task according to Bloom's taxonomy.

Statistically significant results were obtained for all four indicators of coherence with VITAMIN providing an improvement over JIMI for each indicator: (1) the number of correctly answered tasks; (2) the user confidence in correctly answered tasks; (3) the time to answer tasks correctly; and (4) the user satisfaction with the interface.

15. SUBJECT TERMS Remote Evaluation; Mediator; Software Agents; Coherence; Indirect Manipulation; Heterogeneous Legacy Data; Visual Interface; Usability Metrics; IMPACT; HERMES; JIMI; VITAMIN; JUICE; Army War Reserve Prepositioned Stocks

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UNCLASSIFIED					Dr. Rachelle S. Heller
a. REPORT	b. ABSTRACT	c. THIS PAGE		295	19b. TELEPHONE NUMBER *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED			202 994-5906

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

C O P Y R I G H T

© Copyright 2000 by Joseph Hughes Schafer

All Rights Reserved

Sponsored in part by The United States Army.

Unclassified. Distribution Unlimited

A B S T R A C T

Remote Evaluation of the Coherence of Indirect Manipulation Interface Systems for Agent-Mediated Legacy Data

By Joseph Hughes Schafer

Directed by Rachelle Silverman Heller

Many information systems depend heavily on distributed legacy data sources. These data sources introduce a number of significant problems, especially when the sources must be combined and displayed to remote users. Many researchers have investigated various interface systems, however empirical studies have not been published that examine remote interfaces to distributed heterogeneous data. The purpose of this research is to determine the efficacy of a system that provides a more coherent representation of this distributed data in comparison to a more traditional system for users performing representative tasks.

This dissertation presents the results of remote usability experiments in a specific, well-defined context. These web-based experiments empirically determine whether coherence is enhanced through application of the proposed methodology by presenting each interface system and a sequence of representative tasks. The remote evaluation system measures coherence based upon the subject's time to complete each task, the correctness of their answer, and their subjective confidence in that answer. When all tasks have been completed, the respondents complete a usability survey to express their satisfaction with the interface system.

The specific research undertaking is to determine whether a system based upon this proposed methodology, the Visual Interface To Agent Mediated Information Networks (VITAMIN) system, is superior to a system based upon a traditional approach, the Java Indirect Manipulation Interface (JIMI) system. VITAMIN was developed to add

coherence to the legacy data and JIMI was developed as a control treatment to represent a traditional legacy approach. Both VITAMIN and JIMI were implemented as *indirect manipulation* interface systems, or non-anthropomorphic interface agents.

The interface systems provide intermediate query predicate actions to the Heterogeneous Reasoning and Mediator System, HERMES. HERMES is an agent mediator system in the IMPACT agent architecture that provides access to legacy data sources for this research. The remote usability experiment was enabled with the Java Usability Interface Comparison and Evaluation (JUICE) system, which was also developed for this research. JUICE also uses HERMES to chronicle the experimental results.

The experiments used a within-subject randomly counter-balanced design involving 56 subjects. An additional 63 subjects were tested in two phases of pilot experiments. The JUICE system randomly varied the order of the treatments to counterbalance the experiment. The representative tasks were divided into three types. This division was based upon the levels of comprehension, application, and analysis required to answer the task according to Bloom's taxonomy.

Statistically significant results were obtained for all four indicators of coherence with VITAMIN providing an improvement over JIMI for each indicator: (1) the number of correctly answered tasks; (2) the user confidence in correctly answered tasks; (3) the time to answer tasks correctly; and (4) the user satisfaction with the interface.

D E D I C A T I O N

> If many faultes in this book you fynde, Yet think not the correctors blynde; If Argos heere hymselfe had beene, he should perchance not all have seene. – 1565 -Richard *Shacklock [SHA65]*

To my wife, Maureen, and our children: Joseph Palmer, Mary Bridget, Thomas John, Katherine Ann, (James Palmer and John Hughes – RIP), Peter Gabriel, and Lucy Aileena for their unwavering love and support.

To my parents, Joseph Robert and Sarah Jean (Hughes) Schafer, for their love, wisdom, and reminders to "promise to get back to being a good father, husband, brother, and son once this dissertation is complete."

To all of my friends and colleagues and to my teachers and mentors, especially:

Mrs. Rita Braden, High School Chemistry teacher, for my first thorough appreciation and fine immersion in the scientific method.

Mrs. Linda Giffin, Eleventh-Grade English teacher, for tremendous inspiration and support. She guided my first research paper and presentation on Samuel Johnson's Dictionary. She wrote a terrific letter recommending my admission to West Point.

The Honorable William Willoughby, Civilian Assistant to the Secretary of the Army for recommending that Senator John Glenn appoint me to West Point. Bill was the US Military Academy Liaison Officer (MALO) in Cleveland, Ohio who has remained a great friend and mentor.

Colonel Stephen Klinefelter for being my first, finest, and continuing mentor. A VMI and NPS Graduate, “Captain and then Major K” was my sponsor while I was a West Point cadet. Stephen found me as a Plebe living in the computer labs and took me under his wing.

Colonel Andre Sayles, Ph.D. for assisting in the preparation of my first published paper, “Emulating A Vector Associative Processor,” which won third prize in the IEEE student paper contest. Now as Professor and Head of the Department of EE & CS at West Point, he continues to encourage me in many ways.

Colonel Frank Ward for recommending me for command of the Brigade Headquarters Company while he was my Battalion Commander and providing continuing advice.

A C K N O W L E D G M E N T S

The author wishes to thank many people, especially my advisory committee.

Mr. Tim "TJ" Rogers, University of Maryland – College Park (UMCP) programmer extraordinaire, for helping to make the theory work! TJ worked tirelessly to inspire, code, and debug many prototype systems upon which the research systems developed for this work relied.

Professor Rachelle S. Heller, Ph.D., Advocate and Director of Research, Computer Science Department, Interim Associate Dean for Academic Affairs, School of Engineering and Applied Science, GWU, for introducing the fields of educational technology and multimedia and for her support, patience, and constructive criticism.

Professor C. Dianne Martin, Ed.D., Computer Science Department, GWU, for also introducing the fields of educational technology and multimedia and for her constant encouragement and optimism.

Professor Thomas Nagy, Ph.D. Professor, Department of Management Science, School of Business and Public Management, GWU, for reinforcing artificial intelligence, user interface, and statistical analysis concepts.

Professor Bhagirath Narahari, Ph.D. Chair and Professor, Computer Science Department, School of Engineering and Applied Science, GWU, for reinforcing the concepts of automata and formal languages and for advice on distributed databases and information retrieval.

Professor V. S. Subrahamanian, Ph.D. Associate Professor, Computer Science, Department of Computer Science, The University of Maryland, College Park, for

generously sharing his work with mediator systems and agent architectures and for his advice on multimedia data and agents.

Lt. Colonel John A. Marin, Ph.D., (non-voting member) Associate (Academy) Professor, Department of Electrical Engineering and Computer Science; Director of the Information Technology and Operations Center at USMA (formerly the Office of Artificial Intelligence, Analysis, and Evaluation) for allowing flexibility to complete this research.

West Point Cadets, and the USMA Staff and Faculty especially CS489, CS486, and CS408 groups and the subjects from PL100.

GWU professors and fellow students, especially the Interactive MultiMedia and Educational Technology (IMM/ET) and Artificial Intelligence research groups for their support and advice.

Colleagues and friends who helped review this document, especially Dave Nash, Dan Ragsdale, and Scot Ransbottom.

TABLE OF CONTENTS

Title	*Page*
COPYRIGHT	**iii**
ABSTRACT	**iv**
DEDICATION	**vi**
ACKNOWLEDGMENTS	**viii**
TABLE OF CONTENTS	**x**
LIST OF FIGURES	**xvii**
LIST OF TABLES	**xx**
LIST OF SYMBOLS	**xxii**
GLOSSARY OF TERMS	**xxvii**
Chapter 1: INTRODUCTION	**1**
1.1 DOMAIN MOTIVATION	6
1.2 RESEARCH PROBLEM	7
1.3 RESEARCH METHODOLOGY	7
Chapter 2: LITERATURE REVIEW	**11**
2.1 INTERFACE SYSTEM USABILITY	13
2.2 USABILITY METRICS AND REMOTE EVALUATION	17

2.3 COHERENCE 22

2.4 INDIRECT MANIPULATION 28

2.5 HETEROGENEOUS LEGACY DATA SOURCES 35

2.5.1 Army Unit Organization and Data 36

2.5.2 Software Agents 42

2.6 OPEN ISSUES 50

Chapter 3: METHODOLOGY 52

3.1 RESEARCH SYSTEMS AND HYPOTHESES 53

3.1.1 Visual Interface to Agent Mediated Information Networks (VITAMIN) System 54

3.1.2 The Java Indirect Manipulation Interface (JIMI) System 60

3.1.3 Java Usability Interface Comparison and Evaluation (JUICE) System 61

3.1.4 Mediator Agent Actions 63

3.1.5 Development Environment 64

3.1.6 Research Problem 65

3.1.7 Research (Alternative) Hypotheses 65

3.1.8 Testable Null Hypotheses 66

3.2 EXPERIMENTAL DESIGN 67

3.2.1 Independent Variable 67

3.2.2 Dependent Variables 68

3.2.3 Pilot Study Experiments 68

3.2.3.1 Representative Tasks 70

3.2.3.2 Subjects and Questionnaires 70

3.2.4 Dissertation Experiments 71

3.2.4.1 Representative Tasks 73

3.2.4.2 Subjects and Questionnaires 75

3.2.5 Research Procedures 75

3.2.6 Data Collection and Analysis 85

3.2.7 Reliability and Validity 88

3.2.8 Materials 88

3.3 ETHICAL CONSIDERATIONS AND ISSUES 89

3.4 METHODOLOGY SUMMARY 90

Chapter 4: RESULTS AND ANALYSIS 91

4.1 RESEARCH QUESTIONS 92

4.2 QUANTITATIVE ANALYSIS PROCEDURES 93

4.3 NUMBER OF CORRECTLY ANSWERED TASKS 96

4.3.1 Descriptive Statistics - Correctness 96

4.3.2 Hypothesis Test - Correctness 97

4.3.3 Type I Tasks - Correctness 97

4.3.4 Type II Tasks - Correctness 98

4.3.5 Type III Tasks - Correctness 99

4.3.6 Washout Assumption Test - Correctness 101

4.4 USER CONFIDENCE IN CORRECTLY ANSWERED TASKS 102

4.4.1 Descriptive Statistics – Confidence 102

4.4.2 Hypothesis Test - Confidence 103

4.4.3 Type I Tasks - Confidence 104

4.4.4 Type II Tasks - Confidence 105

4.4.5 Type III Tasks - Confidence 106

4.4.6 Washout Assumption Test - Confidence 107

4.5 TIME TO ANSWER TASKS CORRECTLY 108

4.5.1 Descriptive Statistics - Speed 109

4.5.2 Hypothesis Test - Speed 110

4.5.3 Type I Tasks - Speed 110

4.5.4 Type II Tasks - Speed 111

4.5.5 Type III Tasks - Speed 112

4.5.6 Washout Assumption Test - Speed 114

4.6 USER SATISFACTION WITH THE INTERFACE 115

4.6.1 Descriptive Statistics - Satisfaction 115

4.6.2 Hypothesis Test - Satisfaction 116

4.6.3 Washout Assumption Test - Satisfaction 117

4.6.4 Individual Satisfaction Measures 118

4.6.5 Reliability Analysis - Satisfaction 126

4.7 QUALITATIVE RESULTS 128

4.7.1 Correctness 128

4.7.2 Confidence 129

4.7.3 Speed 129

4.7.4 Satisfaction 130

4.7.5 Overall Results 130

4.7.6 Qualitative Analysis 132

4.8 ANALYSIS SUMMARY 133

Chapter 5: CONCLUSION AND FUTURE WORK 138

5.1 CONCLUSIONS AND CONTRIBUTIONS 139

5.2 ADDITIONAL RESEARCH OPPORTUNITIES 145

5.2.1 Future Work in Usability 146

5.2.2 Future Work in Methodology 150

5.3 DISSERTATION SUMMARY 152

REFERENCE LIST 154

Appendix A: SUBJECT INSTRUCTIONS 170

Appendix B: SUBJECT BACKGROUND SURVEY 176

Appendix C: EXPERIMENTAL TASK LISTS 178

Appendix D: USER PREFERENCE QUESTIONNAIRES 183

Appendix E: JUICE SYSTEM EVALUATION SURVEY QUESTIONS 186

Appendix F: HUMAN RESEARCH POLICIES 187

Appendix G: DISSERTATION EXPERIMENT FORMS 193

Appendix H: EXAMPLE OF RAW DATA 197

Appendix I: SOURCE CODE 198

Appendix J: QUANTITATIVE DATA 210

Appendix K: QUALITATIVE DATA **215**

Appendix L: STATISTICS **222**

Appendix M: ETHICAL RESEARCH **248**

LIST OF FIGURES

Number *Page*
Figure 1 JIMI Screen Capture 3
Figure 2 VITAMIN Screen Capture 4
Figure 3 Research Systems Relationships 5
Figure 4 JUICE Task with VITAMIN 8
Figure 5 ISO Usability Factors 13
Figure 6 MUSiC Determinants - Bevan, Kirakowski, and Maissel [BEV91] 14
Figure 7 Usability Factor Relationships - Bevan and Macleod [BEV93] 16
Figure 8 QUIS component groups - LAP [LAP99] 19
Figure 9 SUMI Measurement Scales - HFRG [HFR99] 20
Figure 10 Coherence Metrics 26
Figure 11 Direct Manipulation - Shneiderman [SHN83] 30
Figure 12 Interface Agents - Maes [MAE94a] 32
Figure 13 Indirect Manipulation 34
Figure 14 Army Desert Storm Organization 37
Figure 15 Army Division Organization 38
Figure 16 Separate Brigade Organization - HQDA [HQD96a] 40
Figure 17 Generic Agent Architecture - Huhns and Singh [HUH98a] 43
Figure 18 Levels of IS Interoperability - C4ISR-AWG [C4I98b] 45
Figure 19 Central Mediator Agent - Huhns [HUH98c] 48
Figure 20 VITAMIN Screen Capture 56
Figure 21 Equipment Readiness On Fictitious Supply Ships 57
Figure 22 A Drill-Down Of The Fictitious Supply Ship Alexandria 58
Figure 23 A Drill-Down To The Unit Level 59
Figure 24 JIMI Screen Capture 60
Figure 25 JUICE logon 62
Figure 26 Bloom's taxonomy [BLO56] 74
Figure 27 Java Usability Interface Comparison and Evaluation (JUICE) 76
Figure 28 JUICE logon 77
Figure 29 JUICE Welcome 78
Figure 30 JUICE Demographic Survey 79
Figure 31 JUICE Task with VITAMIN 80
Figure 32 JUICE Task with JIMI 81
Figure 33 VITAMIN / JIMI Comparison 82
Figure 34 JUICE VITAMIN JIMI Comments 83
Figure 35 JUICE Evaluation 84
Figure 36 JUICE Logout 85
Figure 37 Number of Correctly Answered Tasks Box-plot 96
Figure 38 Number of Correctly Answered Tasks – Type I Box-plot 98

Figure 39 Number of Correctly Answered Tasks – Type II Box-plot 99
Figure 40 Number of Correctly Answered Tasks – Type III Box-plot..... 100
Figure 41 Number of Correctly Answered Tasks – Washout Box-plot..... 101
Figure 42 Confidence in Correctly Answered Tasks Box-plot..... 103
Figure 43 Confidence in Correctly Answered Tasks – Type I Box-plot 104
Figure 44 Confidence in Correctly Answered Tasks – Type II Box-plot..... 105
Figure 45 Confidence in Correctly Answered Tasks – Type III Box-plot..... 106
Figure 46 Confidence in Correctly Answered Tasks – Washout Box-plot 108
Figure 47 Time to Correctly Answer Tasks Box-plot..... 109
Figure 48 Time to Correctly Answer Tasks – Type I Box-plot..... 111
Figure 49 Time to Correctly Answer Tasks – Type II Box-plot 112
Figure 50 Time to Correctly Answer Tasks – Type III Box-plot..... 113
Figure 51 Time to Correctly Answer Tasks – Washout Box-plot..... 114
Figure 52 User Satisfaction with the IMIS Box-plot..... 116
Figure 53 User Satisfaction with the IMIS – Washout Box-plot..... 117
Figure 54 The System Was Easy To Use. Box-plot..... 119
Figure 55 The System Was Helpful. Box-plot..... 120
Figure 56 The System Allowed Me To Perform Faster. Box-plot..... 121
Figure 57 The System Provided High Information Quality. Box-plot..... 122
Figure 58 The System Provided High Interface Quality. Box-plot..... 123
Figure 59 The System Allowed Me To Learn About The Data. Box-plot..... 124
Figure 60 The System Was Enjoyable To Use. Box-plot..... 125
Figure 61 The System Was Useful. Box-plot..... 126
Figure 62 Correctness – Type I Washout Box-plot..... 231
Figure 63 Confidence – Type I Washout Box-plot..... 231
Figure 64 Speed – Type I Washout Box-plot..... 231
Figure 65 Correctness – Type II Washout Box-plot 231
Figure 66 Confidence – Type II Washout Box-plot..... 231
Figure 67 Speed – Type II Washout Box-plot..... 231
Figure 68 Correctness – Type III Washout Box-plot..... 232
Figure 69 Confidence – Type III Washout Box-plot 232
Figure 70 Speed – Type III Washout Box-plot..... 232
Figure 71 Correctness – Task 4 Box-plot..... 235
Figure 72 Confidence – Task 4 Box-plot..... 235
Figure 73 Speed – Task 4 Box-plot..... 235
Figure 74 Correctness – Task 5 Box-plot..... 235
Figure 75 Confidence – Task 5 Box-plot..... 235
Figure 76 Speed – Task 5 Box-plot..... 235
Figure 77 Correctness – Task 6 Box-plot..... 236
Figure 78 Confidence – Task 6 Box-plot..... 236
Figure 79 Speed – Task 6 Box-plot..... 236
Figure 80 Correctness – Task 7 Box-plot..... 236
Figure 81 Confidence – Task 7 Box-plot..... 236

Figure 82 Speed – Task 7 Box-plot 236
Figure 83 Correctness – Task 8 Box-plot 237
Figure 84 Confidence – Task 8 Box-plot 237
Figure 85 Speed – Task 8 Box-plot 237
Figure 86 Correctness – Task 9 Box-plot 237
Figure 87 Confidence – Task 9 Box-plot 237
Figure 88 Speed – Task 9 Box-plot 237
Figure 89 Correctness – Task 10 Box-plot 238
Figure 90 Confidence – Task 10 Box-plot 238
Figure 91 Speed – Task 10 Box-plot 238
Figure 92 Correctness – Task 11 Box-plot 238
Figure 93 Confidence – Task 11 Box-plot 238
Figure 94 Speed – Task 11 Box-plot 238
Figure 95 Correctness – Task 12 Box-plot 239
Figure 96 Confidence – Task 12 Box-plot 239
Figure 97 Speed – Task 12 Box-plot 239
Figure 98 Correctness – Task 4 Washout Box-plot 242
Figure 99 Confidence – Task 4 Washout Box-plot 242
Figure 100 Speed – Task 4 Washout Box-plot 242
Figure 101 Correctness – Task 5 Washout Box-plot 242
Figure 102 Confidence – Task 5 Washout Box-plot 242
Figure 103 Speed – Task 5 Washout Box-plot 242
Figure 104 Correctness – Task 6 Washout Box-plot 243
Figure 105 Confidence – Task 6 Washout Box-plot 243
Figure 106 Speed – Task 6 Washout Box-plot 243
Figure 107 Correctness – Task 7 Washout Box-plot 243
Figure 108 Confidence – Task 7 Washout Box-plot 243
Figure 109 Speed – Task 7 Washout Box-plot 243
Figure 110 Correctness – Task 8 Washout Box-plot 244
Figure 111 Confidence – Task 8 Washout Box-plot 244
Figure 112 Speed – Task 8 Washout Box-plot 244
Figure 113 Correctness – Task 9 Washout Box-plot 244
Figure 114 Confidence – Task 9 Washout Box-plot 244
Figure 115 Speed – Task 9 Washout Box-plot 244
Figure 116 Correctness – Task 10 Washout Box-plot 245
Figure 117 Confidence – Task 10 Washout Box-plot 245
Figure 118 Speed – Task 10 Washout Box-plot 245
Figure 119 Correctness – Task 11 Washout Box-plot 245
Figure 120 Confidence – Task 11 Washout Box-plot 245
Figure 121 Speed – Task 11 Washout Box-plot 245
Figure 122 Correctness – Task 12 Washout Box-plot 246
Figure 123 Confidence – Task 12 Washout Box-plot 246
Figure 124 Speed – Task 12 Washout Box-plot 246

LIST OF TABLES

Number *Page*
Table 1 Data Sources 63
Table 2 Pilot Research Notation 69
Table 3 Pilot Research Design Structure 69
Table 4 Dissertation Research Notation 72
Table 5 Dissertation Experiment Research Design Structure 72
Table 6 Task-Type Characteristics 73
Table 7 Research Questions 92
Table 8 Number of Correctly Answered Tasks 96
Table 9 Number of Correctly Answered Tasks – Type I 97
Table 10 Number of Correctly Answered Tasks – Type II 98
Table 11 Number of Correctly Answered Tasks – Type III 99
Table 12 Number of Correctly Answered Tasks – Washout 101
Table 13 Confidence Analysis Values 102
Table 14 Confidence in Correctly Answered Tasks 102
Table 15 Confidence in Correctly Answered Tasks – Type I 104
Table 16 Confidence in Correctly Answered Tasks – Type II 105
Table 17 Confidence in Correctly Answered Tasks – Type III 106
Table 18 Confidence in Correctly Answered Tasks – Washout 107
Table 19 Time to Correctly Answer Tasks 109
Table 20 Time to Correctly Answer Tasks – Type I 110
Table 21 Time to Correctly Answer Tasks – Type II 111
Table 22 Time to Correctly Answer Tasks – Type III 112
Table 23 Time to Correctly Answer Tasks – Washout 114
Table 24 Survey Analysis Values 115
Table 25 User Satisfaction with the IMIS 115
Table 26 User Satisfaction with the IMIS – Washout 117
Table 27 User Satisfaction Individual Measures - Summary 118
Table 28 User Satisfaction Individual Measures - Hypotheses 118
Table 29 Discriminatory Analysis 127
Table 30 Overall Results 134
Table 31 Results by Task-Type 135
Table 32 Results by Question 136
Table 33 Confidence Values 178
Table 34 Ship Name Mappings 182
Table 35 Survey Analysis Scale 185
Table 36 Variables 211
Table 37 Subject Task Responses 212
Table 38 Subject Totals and Correct Task Averages 213

Table 39 Subject Satisfaction Survey Responses 214
Table 40 Raw Qualitative Comments 221
Table 41 Descriptive Statistics – All Questions 223
Table 42 T-Test – All Questions 223
Table 43 Descriptive Statistics – All Questions - Washout 224
Table 44 T-Test – All Questions - Washout 225
Table 45 Descriptive Statistics – Satisfaction Survey 226
Table 46 T-Test – Satisfaction Survey 227
Table 47 Descriptive Statistics – By Type 228
Table 48 T-Test – By Type 229
Table 49 Descriptive Statistics – By Type – Washout 230
Table 50 T-Test – By Type - Washout 233
Table 51 Descriptive Statistics – By Question 234
Table 52 T-Test – By Question 240
Table 53 Descriptive Statistics – By Question – Washout 241
Table 54 T-Test – By Question - Washout 247

LIST OF SYMBOLS

This section is a quick reference to acronyms which readers may find unfamiliar. Several acronym and glossary entries were inspired by DISA [DIS98b] and Howe [HOW98].

ABE	Agent Building Environment
AIC	Artificial Intelligence Center (US Army – now ASACC)
API	Application Program Interface
APS	Army Prepositioned Stocks
ARL	Army Research Laboratory
ARO	Army Research Office
ASACC	Army Strategic and Advanced Computing Center (formerly the Artificial Intelligence Center)
ATA	Army Technical Architecture
AWR	Army War Reserve
BDE	Brigade
BN	Battalion
BNTF	Battalion Task Force
C2	Command and Control
C^2S^2	Correctly, Confidently, Speedily, Satisfactorily
C3I	Command, Control, Communications, and Intelligence
C4I	Command, Control, Communications, Computers, and Intelligence
C4ISR	Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance
CATS	Collaborative Agent Technology System
CBD	Counterbalanced Design
CG	Commanding General
CHIMP	Collaborative Heterogeneous Interactive Multimedia Platform
CICS	Customer Information Control System
CINC	Commander In Chief
CLXXI	Classroom twenty-one
COE	Common Operating Environment
COM	Common Object Model
CONUS	Continental United States
CORBA	Common Object Request Broker Architecture
COSE	Common Open Software Environment
COTS	Commercial Off-the-Shelf
CS	Combat Support; Computer Science

CSCW	Computer Supported Cooperative Work
CSS	Combat Service Support
D.Sc.	Doctor of Science, also Sc.D.
D/EECS	Department of Electrical Engineering and Computer Science
DARPA	Defense Advanced Research Projects Agency
DB	Database
DBMS	Database Management System
DCOM	Distributed Component Object Mode
DII	Defense Information Infrastructure
DIS	Distributed Interactive Simulation
DLA	Defense Logistics Agency
DMSO	Defense Modeling and Simulation Office
DoD; DOD	Department of Defense
DQI	Dynamic Query Interface
DUIC	Derivative UIC (Unit Identification Code)
DW	Date Warehouse
EAC	Echelons Above Corps
EAD	Echelons Above Division
EECS	Electrical Engineering and Computer Science
EquipRU	Equipment Roll Up
ERC	Equipment Requirements Code
GCCS	Global Command and Control System
GCSS	Global Combat Support System
GIS	Geographic Information System
GOMS	Goals, Operators, Methods, and Selection Rules
GOTS	Government Off-the-Shelf
GUI	Graphical User Interface
HCI	Human-Computer Interface
HERMES	Heterogeneous Reasoning and Mediator System
HLA	High Level Architecture
IIOP	Internet Inter-Orb Protocol
IMIS	Indirect Manipulation Interface System
IMPACT	Interactive Maryland Platform for Agents Collaborating Together
IS	Information System
ISO	International Organization for Standardization (not acronym, prefix: *-iso* = “same” in *Greek*)
ISR	Intelligence, Surveillance & Reconnaissance
IT	Information Technology
ITOC	Information Technology and Operations Center
ITW	Information Technology Warrior | Warfare
IW	Information Warfare

JANUS	Janus
JIMI	Java Indirect Manipulation Interface
JKQML	Java Knowledge Query and Manipulation Language
JNI	Java Native Interface
JSA	Joint Systems Architecture
JTA	Joint Technical Architecture
JUICE	Java Usability Interface Comparison & Evaluation tool
JV	Joint Vision
JVM	Java Virtual Machine
KQML	Knowledge Query and Manipulation Language
LIA	Logistics Integration Agency (US Army)
LIN	Line Item Number
LISI	Levels of Information Systems Interoperability
LOC	Location
LOGTAADS	Logistics TAADS
M&S	Modeling and Simulation
MAS	Mobile Agent System
MIDB	Management Information Database
MIS	Management Information Systems
MIS	Multimedia Information Systems
MIT	Massachusetts Institute of Technology
MOA	Memorandum of Agreement
ModSAF	Modular Semi-Automated Forces
MOM	Message-oriented middleware
MS	Microsoft
NASA	National Aeronautics and Space Administration
NCSC	National Computer Security Center
NIMA	National Imagery and Mapping Agency
NIPNET or NIPRNET	Non-secure IP Routing Network
NIST	National Institute of Standards and Technology
NRO	National Reconnaissance Office
NSA	National Security Agency
NT	New Technology
OAIAE	Office of Artificial Intelligence, Analysis, and Evaluation (USMA) – ITOC as of June 1999
ODBC	Open Database Connectivity
ODMG	Object Data Management Group
OLAP	On-Line Analytical Processing
OLE	Object Linking and Embedding
OMG	Object Management Group
OOTW	Operations Other Than War

ORB	Object Request Broker
OS	Operating System
OSD	Office of the Secretary of Defense
OSD A&T	Office of the Secretary of Defense for Acquisition and Technology
OSF	Open Software Foundation
OSI	Open Systems Interconnection
OTP	On-Line Transaction Processing
PM	Program Manager
RDBMS	Relational Database Management System
RMI	Remote Method Invocation (Java)
RMON	Remote Monitoring
RPC	Remote Procedure Call
RTI	Run Time Infrastructure
RU	Roll-Up
SACC	(Army) Strategic and Advanced Computing Center. also ASACC
SDF	Simulation Data Format
SE	Synthetic Environments; Systems Engineering
SEDRIS	Synthetic Environment Data Representation and Interchange Specification
SFA	Sales Force Automation
SGI	Silicon Graphics, Inc.
SGML	Standard Generalized Markup Language
SHADE	SHAred Data Environment
SIPRNET	Secure IP Routing Network
SOM	Simulation Object Model
SQL	Structured Query Language
SSL	Secure Socket Layer
STD	Standard
SUN	Stanford University Networks
SW	Software
SWE	Software Engineering
TAADS	The Army Authorization Documents System
TAFIM	Technical Architecture Framework for Information Management
TAM	Technology Acceptance Model
TCL	Tool Command Language
TCP/IP	Transmission Control Protocol / Internet Protocol
TCSEC	Trusted Computer Security Evaluation Criteria
TF	Task Force
TFXXI	Task Force Twenty-One

TP	Transport Protocol
TP	Transaction Processing
TSIMMIS	The Stanford-IBM Manager of Multiple Information Sources
UI	User Interface
UIC	Unit Identification Code
UML	Unified Modeling Language
VITAMIN	Visual Interface to Mediated Information Networks
VRML	Virtual Reality Modeling Language
WIMP	Windows, Icons, Menus, Pointing Device (mouse)
XML	eXtensible Markup Language
yp	Yellow Pages

GLOSSARY OF TERMS

> It often does more harm than good to force definitions on things we don't understand. Besides, only in logic and mathematics do definitions ever capture concepts perfectly. The things we deal with in practical life are usually too complicated to be represented by neat, compact expressions. Especially when it comes to understanding minds, we still know so little that we can't be sure our ideas about psychology are even aimed in the right directions. In any case, one must not mistake defining things for knowing what they are. - *Minsky [MIN85]*

Definitions adopted by researchers are often not uniform. A glossary and a list of symbols are provided to establish positions taken in this thesis and to assist with controversial or domain-specific terms.

4GL, fourth generation language
An application-specific, often non-procedural, high level programming language. Martin [MAR85].

Access Control
Process of limiting access to the resources of an IT product only to authorized users, programs, processes, systems, or other IT products.

Access Transparency
Masks differences in data representation and invocation mechanisms to enable interworking between objects. This transparency solves many of the problems of interworking between heterogeneous systems. RM-ODP [RMO99].

Agent
A process that communicates with and performs information preparation and exchange on behalf of a client or server. This communication may be with another agent and the process may be automatic. Software agents are commonly described as: autonomous, situated, goal-oriented, collaborative, flexible, self-starting, adaptive, sociable, mobile.

An agency relationship exists when a principal party depends upon another party (the agent) to undertake some task on the principal's behalf. - Eisenhardt [EIS89]; Sycara [SYC98a].

"The notion of an agent is meant to be a tool for analyzing systems, not an absolute characterization that divides the world into agents and non-agents" - Russell and Norvig [RUS95].

Application Program Interface (API)

1) The interface, or set of functions, between the application software and the application platform. - NIST [NIS96].
2) The means by which an application designer enters and retrieves information. DISA [DIS96b]; [DIS96c].

Architecture

Orderly arrangement of parts; structure: - AHD [AHD92]

Design of the way components fit together. Particularly of processors, both individual and in general. Also any complex system, e.g. "software architecture", "network architecture". - Howe [HOW98]

The design of application software incorporating protocols and the means for expansion and interfacing with other programs. - Microsoft [MIC97a]

Artificial Intelligence (AI)

The field of computer science concerned with the concepts and methods of symbolic inference by computer and symbolic knowledge representation for use in making inferences. AI can be seen as an attempt to model aspects of human thought using computers. It is also sometimes defined as trying to solve by computer any problem that a human can (at present) solve better.

Authentication

1) To verify the identity of a user, device, or other entity in a computer system, often as a prerequisite to allowing access to resources in a system.
2) To verify the integrity of data that have been stored, transmitted, or otherwise exposed to possible unauthorized modification.

Collaborative Agent Technology System

The CATS project at the University of Maryland is funded by the Army Research Laboratory in Adelphi, Maryland, and seeks to answer the following questions:

- What is an agent?

- What are the requirements for a software platform supporting multiple such agents?
- How will such a software platform be built?

Interactive Maryland Platform for Agents Collaborating Together (IMPACT) reflects the current multi-agent platform.

Character-based interface
A non-bitmapped user interface in which the primary form of interaction between the user and system is through text.

Coherence
Defined by this research as a concise usability metric for statistical analysis. It is particularly appropriate for remote evaluations and includes the coherence factors of tasks performed *correctly, confidently, speedily, and satisfactorily* (C^2S^2).

Denotes consistent similarity between items; a logical, orderly, and aesthetically consistent relationship of parts. Describes the extent to which items are locally constant and properties that make an object coherent may be measured, according to Gröller [GRO92]; Gröller and Purgathofer [GRO95]; Murphy and Medin [MUR92]. Improves the usability and usefulness of an interface system by enabling users to perform representative tasks correctly, confidently, speedily, and satisfactorily (C^2S^2).

Command and Control (C2)
The exercise of authority and direction by a properly designated commander over assigned and attached forces in the accomplishment of the mission. Command and control functions are performed through an arrangement of personnel, equipment, communications, facilities, and procedures employed by a commander in planning, directing, coordinating, and controlling forces and operations in the accomplishment of the mission. - JCS [JCS98a].

Command, Control, Communications, and Computer (C4) Systems
Integrated systems of doctrine, procedures, organizational structures, personnel, equipment, facilities, and communications designed to support a commander's exercise of command and control across the range of military operations. - JCS [JCS98a].

Data Dictionary
A specialized type of database containing metadata that is managed by a data dictionary system; a repository of information describing the characteristics of data used to design, monitor, document, protect, and control data in information systems and databases; an application of a data dictionary system. - DISA [DIS98a].

Data Integrity
1) The state that exists when computerized data is the same as that in the source documents and has not been exposed to accidental or malicious alteration or destruction.
2) The property that data has not been exposed to accidental or malicious alteration or destruction.

Data Model
In a database, the user's logical view of the data in contrast to the physically stored data, or storage structures. A description of the organization of data in a manner that reflects the information structure of an enterprise. - DISA [DIS98a].

Data Warehouse
A database containing recent snapshots of enterprise data. Analysts use this database without slowing day-to-day operations of the production transaction databases. A data warehouse often includes sophisticated indexing to enable OLAP. - Bontempo and Zagelow [BON98c]; Chaudhuri and Dayal [CHA97a]; Colliat [COL96]; Friedland [FRI98a]; Gardner [GAR98].

Defensive Information Operations
The main objective of defensive IO is to help protect and defend information and information systems. Defensive IO integrates and coordinates policies and procedures, operations, personnel, and technology to protect and defend information and information systems. Defensive IO are conducted through information assurance (IA), operations security (OPSEC), physical security, counterdeception, counterpropaganda, counterintelligence, EW, and SIO. Defensive IO ensure timely, accurate, and relevant information access while denying adversaries the opportunity to exploit friendly information and information systems for their own purposes. Offensive IO also can support defensive IO. - HQDA [HQD96c]; JCS [JCS98b].

Direct Manipulation
Human-computer interaction characterized by the following attributes:

- Enables visibility of the object of interest
- Allows rapid, reversible, incremental actions
- Immediate interaction with the object of interest replaces complex command language. - Shneiderman [SHN83]; [SHN97b].

Distributed Interactive Simulation (DIS)
Program to electronically link organizations operating in the following four domains: advanced concepts and requirements; military operations; research, development, and acquisition; and training. (2) A synthetic environment within which humans may interact through simulation(s) at multiple sites networked using compliant architecture, modeling, protocols, standards, and databases. - DMSO [DMS94]; [DMS95b]; [DMS98]

Domain
A distinct functional area that can be supported by a family of systems with similar requirements and capabilities. An area of common operational and functional requirements.

Efficacy
Power of efficiency and effectiveness.

Empirical
Work derived from, verifiable by, or relying on experiment or observation, often guided by practical experience more than abstract theory. Experiences are usually descriptive in nature while experiments address cause and effect relationships.

Extensible Markup Language (XML)
XML describes a class of data objects called XML documents and the behavior of computer programs that process them. By construction, XML documents are conforming SGML documents. XML documents are made up of storage units called entities, which contain either parsed or unparsed data. Parsed data is made up of characters, some of which form character data, and some of which form markup. Markup encodes a description of the document's storage layout and logical structure. XML provides a mechanism to impose constraints on the storage layout and logical structure. - Bray, Paoli, and Sperberg-McQueen [BRA97] Harold [HAR99].

Federate
A member of an HLA Federation. All applications participating in a Federation are called Federates. In reality, this may include Federate Managers, data collectors, live entity surrogates, simulations, or passive viewers. - DMSO [DMS97]

Federation
A named set of interacting federates, a common federation object model, and supporting RTI, that are used as a whole to achieve some specific objective. - DMSO [DMS97]

Federation Object Model (FOM)
An identification of the essential classes of objects, object attributes, and object interactions that are supported by an HLA federation. In addition, optional classes of additional information may also be specified to achieve a more complete description of the federation structure and/or behavior. - DMSO [DMS97]

Framework
A fundamental structure, as for a written work or a system of ideas - AHD [AHD92].

In object-oriented programming, a reusable basic design structure, consisting of abstract and concrete classes that assist in building applications - Microsoft [MIC97b].

In object-oriented systems, a set of classes that embodies an abstract design for solutions to a number of related problems - Howe [HOW98].

Goals, Operators, Methods, and Selection Rules (GOMS)
A framework for analyzing routine human computer interaction. - Card, Moran, and Newell [CAR83].

Graphical User Interface (GUI)
System component that allows the user to effect commands, enter transaction sequences, and receive displayed information through graphical representations of objects (menus, screens, buttons, etc.).

HERMES (HEterogeneous Reasoning and MEdiator System)
HERMES is a system for semantically integrating distinct and possibly heterogeneous information sources and reasoning systems. This is accomplished by executing programs, called *mediators*, written in the HERMES system. Mediators are guidelines of how information from different sources will be combined and integrated. The HERMES system is based on the theory of Hybrid Knowledge Bases. In this framework, external information sources are abstracted as domains, which execute certain functions with pre-specified input and output types. These domains are accessed by mediators using an annotated logic-based declarative language. The language provides a powerful and extensible programming environment. The system also provides a uniform environment for the addition of new external sources to existing mediators. - Candan, Subrahmanian, and Rangan [CAN96f]; Subrahmanian et al. [SUB97a].

High Level Architecture (HLA)
Major functional elements, interfaces, and design rules, pertaining as feasible to all DoD simulation applications, and providing a common framework within which specific system architectures can be defined. - DMSO [DMS97]

Human-Computer Interface (HCI)
Hardware and software allowing information exchange between the user and the computer. Also Human-Computer Interaction.

Hybrid Graphical User Interface
A GUI that is composed of tool kit components from more than one user interface style.

HyperText Markup Language (HTML)
The publishing language of the World Wide Web that has text, multimedia, and hyperlink features. A subset of SGML. - Raggett, Hors, and Jacobs [RAG99]

Indirect Manipulation (IM)
Human-computer interaction metaphor defined in this dissertation characterized by the following attributes:

- Incomplete visibility of objects and actions
- Distributed data
- Delayed actions
- Remote devices and actions
- Delayed or incomplete feedback

Indirect manipulation lacks several attributes required by direct manipulation such as immediate, reversible actions. IM represents a compromise between direct manipulation and interface agents.

Information Assurance (IA)
An element of Defensive Information Operations. IA protects and defends information and information systems by ensuring their availability, integrity, identification and authentication, confidentiality, and non-repudiation. - JCS [JCS98b]; Waltz [WAL98].

Information Operations (IO)
Actions taken to affect adversary information and information systems while defending one's own information and information systems. Offensive and defensive Information Operations target information or information systems in order to affect the information-based process, whether human or automated. - JCS [JCS98b]

Information Efficacy (IE)
The ability to use information provided by a system. Denotes the usability and usefulness of an interface system. Depicted as a foundation for information operations, which adds the capabilities of information awareness, decision support, and command and control to the activities of traditional information operations.

Information Technology (IT)
Computers, ancillary equipment, software, firmware and similar procedures, services (including support services), and related resources or any equipment or interconnected system or subsystem of equipment, that is used in the automatic acquisition, storage, manipulation, management, movement, control, display, switching, interchange, transmission, or reception of data or information by the executive agency. - ITMRA [ITM96].

Information Technology and Operations (ITO)
Information operations and all militarily significant information applications, especially their protection and exploitation.

Information Warfare (IW)
Information operations during wartime. - JCS [JCS98b]

Intelligence
The ability of an entity to synthesize an appropriate response that is correlated with a stimulus. Knowledge is the instantiation of intelligence. - Bock [BOC93].

Military Intelligence:

1) The product resulting from the collection, processing, integration, analysis, evaluation, and interpretation of available information concerning foreign countries or areas.
2) Information and knowledge about an adversary obtained through observation, investigation, analysis, or understanding. - JCS [JCS98a].

Interactive Maryland Platform for Agents Collaborating Together (IMPACT)
IMPACT provides a set of servers (yellow pages, thesaurus, registration, type and interface) that facilitate agent interoperability in an application independent manner. In IMPACT, agents have an associated set of service descriptions, specifying the services that they provide. An HTML-like language is used to describe these services. When an agent wishes to identify another agent that provides some service, the requested service must be matched, using a metric, against existing service descriptions. A formal framework for finding the closest matches has been developed. Algorithms that compute the *k* nearest matches as well as all matches within a given distance have been developed and implemented. If a match otherwise fails, a thesaurus is used to extend the search through the use of synonyms. - Arisha et al. [ARI98].

Interoperability

1) The ability of two or more systems or components to exchange data and use information. - IEEE [IEE90].
2) The ability of two or more systems to exchange information and to mutually use the information that has been exchanged. - ASB [ASB98].

Interworking

The exchange of meaningful information between computing elements (semantic integration), as opposed to interoperability, which provides syntactic integration among computing elements.

KQML -- Knowledge Query and Manipulation Language

KQML is a language and protocol for exchanging information and knowledge. - KQML [KQM98]

Legacy Environments

Legacy environments could be called legacy architectures or infrastructures and as a minimum consist of a hardware platform and an operating system. Legacy environments are identified for phase-out, upgrade, or replacement. - DISA [DIS96b].

Legacy Systems

Systems that are candidates for phase-out, upgrade, or replacement. Generally, such systems are in this category because they do not comply with data standards or other standards. Legacy system workloads must be converted, transitioned, or phased out (eliminated). Such systems may or may not operate in a legacy environment. - DISA [DIS96b].

Operationally, legacy systems run on hardware and software that is not part of an organization's current strategy. Organizationally, they are "old, inflexible, expensive, non-portable, and undocumented, but indispensable because they support core business functions." - Alderson and Shah [ALD99]

Live, Virtual, and Constructive Simulation

The categorization of simulation into live, virtual, and constructive is problematic, because there is no clear division between these categories. The degree of human participation in the simulation is infinitely variable, as is the degree of equipment realism. This categorization of simulations also suffers by excluding a category for simulated people working real equipment (e.g., smart vehicles). - DMSO [DMS94]; [DMS95b]; [DMS98]

1) <u>Live Simulation</u>. A simulation involving real people operating real systems.

2) Virtual Simulation. A simulation involving real people operating simulated systems. Virtual simulations inject human-in-the-loop (HITL) in a central role by exercising motor control skills (e.g., flying an airplane), decision skills (e.g., committing fire control resources to action), or communication skills (e.g., as members of a C4I team).
3) Constructive Model or Simulation. Models and simulations that involve simulated people operating simulated systems. Real people stimulate (make inputs to) such simulations, but are not directly determine the outcomes.

Location Transparency
Masks the use of information about location in space when identifying and binding to interfaces. This transparency provides a logical view independent of actual physical location. - RM-ODP [RMO99].

Meliorate
To improve. [AHD92]

Message-Oriented Middleware (MOM)

Allows applications on different computing platforms and networks to exchange data reliably and securely by sending data sets to message queues. The queues hold the message until another application retrieves the data. A direct connection between the applications is not required, and the sender needs no information about the retrieving application. - Ouellette [OUE98].

Metadata
Information describing the characteristics of data; data or information about data; descriptive information about an organization's data, data activities, systems, and holdings. - DISA [DIS98a]. A Data Dictionary contains the metadata.

Methodology
Methodology is a body of practices, procedures, and rules used by those who work in a discipline or engage in an inquiry; a set of working methods and the study or theoretical analysis of such working methods. - AHD [AHD92]

A documented set of procedures and guidelines for one or more phases of the software life cycle, such as analysis or design. Many methodologies include a diagramming notation for documenting the results of the procedure; a step-by-step "cookbook" approach for carrying out the procedure; and an objective (ideally quantified) set of criteria for determining whether the results of the procedure are of acceptable quality. - Howe [HOW98]

Middleware
Technically:

- Software between client applications and server control processes (OS, Network Control Programs, and DBMS)
- Single Application Programming Interface (API) for many applications
- Software Services Repository to allow applications to access all middleware environments

Functionally:

- Software that connects functional area applications. - Colonna-Romano and Srite [COL95]; Thompson [THO97]

Model
A physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process. - DMSO [DMS94]; [DMS95b]; [DMS98]

Modeling and Simulation (M&S)
The use of models, including emulators, prototypes, simulators, and stimulators, either statically or over time, to develop data as a basis for making managerial or technical decisions. The terms "modeling" and "simulation" are often used interchangeably. - DMSO [DMS98].

Multimedia (MM)
The presentation of information on a computer using sound, graphics, animation, and text; using various input and output devices.

Object Model
A specification of the objects intrinsic to a given system, including a description of the object characteristics (attributes) and a description of the static and dynamic relationships that exist between objects. - DMSO [DMS97]

Offensive Information Operations
Offensive IO involve the integrated use of assigned and supporting capabilities and activities, mutually supported by intelligence, to affect adversary decision makers and achieve or promote specific objectives. These assigned and supporting capabilities and activities include, but are not limited to, operations security (OPSEC), military deception, psychological operations (PSYOPS), electronic warfare (EW), physical attack/destruction, and special information operations (SIO), and may include computer network attack (CNA). - *JCS [JCS98b]*

On-Line Analytical Processing
Database software that provides an interface for users to transform raw data and interactively examine the results in various dimensions to look for patterns. OLAP usually involves large amounts of diverse data aggregated into a data warehouse. - Chaudhuri and Dayal [CHA97b]; Navin [NAV98].

Ontology
In philosophy, the branch of metaphysics that studies the nature of being and strives for a systematic account of existence. In artificial intelligence, an explicit formal specification of how to represent the objects, concepts, other entities, and relationships in some domain, often expressed as a hierarchical structuring of knowledge about things. A set of agents that share the same ontology will be able to communicate about a domain of discourse without necessarily operating on a globally shared view or theory. - Luke, Spector, and Rager. [LUK96]; Luke et al. [LUK97b]; Taylor, Stoffel, and Hendler [TAY97].

Open System
A system that implements sufficient open specifications for interfaces, services, and supporting formats. An Open System enables properly engineered components to be utilized across a wide range of systems with minimal changes, to interoperate with other components on local and remote systems, and to interact with users in a style that facilitates portability. An open system is characterized by the following:

Well-defined, widely used, non-proprietary interfaces/protocols

Use of standards that are developed/adopted by industrially recognized standards bodies

Definition of all aspects of system interfaces to facilitate new or additional systems capabilities for a wide range of applications

Explicit provision for expansion or upgrading through the incorporation of additional or higher performance elements with minimal impact on the system

- OS-JTF [OSJ98].

Open Systems Approach
An open systems approach is a business approach that emphasizes commercially supported practices, products, specifications, and standards. The approach defines, documents, and maintains a system technical architecture that depicts the lowest level of system configuration control. This architecture clearly identifies all the performance characteristics of the system including those that will be accomplished with an implementation that references open standards and specifications. - OS-JTF [OSJ98].

Open Systems Interconnect (OSI)
(OSI, seven layer model) The OSI Reference Model of network architecture and an implementing suite of protocols (protocol stack) were developed by ISO in 1978 as a framework for international standards in heterogeneous computer network architectures. The architecture is split between seven layers, from lowest to highest: 1 physical layer, 2 data link layer, 3 network layer, 4 transport layer, 5 session layer, 6 presentation layer, 7 application layer. Each layer uses the layer immediately below it and provides a service to the layer above.

Operational Architecture (OA)
An Operational Architecture is a description (often graphical) of the operational elements, assigned tasks, and information flows required to support the warfighter. It defines the type of information, the frequency of the exchange, and what tasks are supported by these information exchanges. - DISA [DIS96b].

Portability
The ease with which a system, component, body of data, or user can be transferred from one hardware or software environment to another. - DISA [DIS96b]; [DIS96c].

Practice
A recommended implementation or process that further clarifies the implementation of a standard or a profile of a standard. - NIMA [NIM98].

Reconnaissance
A mission undertaken to obtain, by visual observation or other detection methods, information about the activities and resources of an enemy or potential enemy, or to secure data concerning the meteorological, hydrographic, or geographic characteristics of a particular area. - JCS [JCS98a].

Reference Model
A reference model is a generally accepted abstract representation that allows users to focus on establishing definitions, building common understandings, and identifying issues for resolution. For Warfare and Warfare Support System (WWSS) acquisitions, a reference model is necessary to establish a context for understanding how the disparate technologies and standards required to implement WWSS relate to each other. Reference models provide a mechanism for identifying essential issues associated with portability, scalability, and interoperability. Most importantly, reference models will aid in the evaluation and analysis of domain specific architectures. - OS-JTF [OSJ98].

Runtime Infrastructure (RTI)
The general purpose distributed operating system software that provides the common interface services during the runtime of an HLA federation. - DMSO [DMS97]

Scalability, (Scaleability)
1) The capability to adapt hardware or software to accommodate changing workloads. - OS-JTF [OSJ98].
2) The ability to use the same application software on many different classes of hardware/software platforms from personal computers to supercomputers (extends the portability concept). The ability to grow to accommodate increased workloads. - DISA [DIS96b]; [DIS96c].

Secondary Imagery Dissemination (SID)
The process for the post-collection electronic transmission of or receipt of C3I exploited non-original imagery and imagery products in other than real or near-real time.

Security
1) The combination of confidentiality, integrity, and availability.
2) The quality or state of being protected from uncontrolled losses or effects. Note: Absolute security may in practice be impossible to reach; thus the security "quality" could be relative. Within state models of security systems,

security is a specific "state" that is to be preserved under various operations. - DISA [DIS96a]; HQDA [HQD98a]; Hu [HUY98a].

Service Area
A set of capabilities grouped into categories by function. The JTA defines a set of services common to DoD information systems. - DISA [DIS98b].

Simulation Object Model (SOM)
A specification of the intrinsic capabilities that an individual simulation offers to federations. The standard format in which SOMs are expressed provides a means for federation developers to quickly determine the suitability of simulation systems to assume specific roles within a federation. - DMSO [DMS97]

Specification
A document prepared to support acquisition that describes the essential technical requirements for purchased materiel and the criteria for determining whether those requirements are met. - DOD [DOD93a].

Standard
A document that establishes uniform engineering and technical requirements for processes, procedures, practices, and methods. Standards may also establish requirements for selection, application, and design criteria of material. - DOD [DOD93a].

Standard Generalized Markup Language (SGML)
ISO 8879 specifies SGML. HTML and XML are forms of SGML.

Standard Simulator Database Interchange Format (SIF)
A DoD data exchange standard, DOD [DOD93b], adopted as an input/output vehicle for sharing externally created simulator databases among the operational system training and mission rehearsal communities.

Surveillance
The systematic observation of aerospace, surface or subsurface areas, places, persons, or things, by visual, aural, electronic, photographic, or other means. - JCS [JCS98a].

System
1) People, machines, and methods organized to accomplish a set of specific functions. - FIPS [FIP91].

2) An integrated composite of people, products, and processes that provide a capability or satisfy a stated need or objective. - DOD [DOD96].

Systems Architecture (SA)
A description, including graphics, of the systems and interconnections providing for or supporting a warfighting function. The SA defines the physical connection, location, and identification of the key nodes, circuits, networks, warfighting platforms, etc., and allocates system and component performance parameters. It is constructed to satisfy Operational Architecture requirements in the standards defined in the Technical Architecture. The SA shows how multiple systems within a domain or an operational scenario link and interoperate, and may describe the internal construction or operations of particular systems in the SA. - DISA [DIS98b].

Task Organization
Task organization is the distribution of assets to subordinate control headquarters under the appropriate command or support relationship - HQDA [HQD92]. The higher headquarters allocates resources to the subordinate, as needed to accomplish the assigned mission, based on the higher headquarters commander's estimate. Assets are not distributed on a "fair share" basis. Subordinate commanders may request more assets from higher headquarters. Task organization is changed during the operation only if changing conditions dictate - HQDA [HQD92].

Technology Acceptance Model (TAM)
A framework for empirically testing user acceptance of new information systems. - Davis, Bagozzi, and Warshaw [DAV89]; Davis [DAV93]; Davis and Venkatesh [DAV95].

Technical Architecture (TA)
The minimal set of rules governing the arrangement, interaction, and interdependence of the parts or elements whose purpose is to ensure that a conformant system satisfies a specified set of requirements. The technical architecture identifies the services, interfaces, standards, and their relationships. It provides the technical guidelines for implementation of systems upon which engineering specifications are based, common building blocks are built, and product lines are developed. - DISA [DIS98b].

Technical Reference Model (TRM)
A conceptual framework that provides the following:

1) A consistent set of service and interface categories and relationships used to address interoperability and open system issues.
2) Conceptual entities that establish a common vocabulary to better describe, compare, and contrast systems and components.
3) A basis (an aid) for the identification, comparison, and selection of existing and emerging standards and their relationships.

The TRM framework is not an architecture. It is not a set of standards nor does it contain standards.

t-Test.
Returns the probability associated with a Student's t-Test. Commonly used to determine whether two samples are likely to have come from the same two underlying populations.

Video
Electro-optical imaging sensors and systems that generate sequential or continuous streaming imagery at specified rates. Video standards are developed by recognized bodies such as ISO, ITU, SMPTE, EBU, etc. - NIMA [NIM98].

Weapon Systems
A combination of one or more weapons with all related equipment, materials, services, personnel and means of delivery and deployment (if applicable) required for self sufficiency. - JCS [JCS98a].

CHAPTER 1: INTRODUCTION

> A man ceases to be a beginner in any given science and becomes a master in that science when he has learned that . . . he is going to be a beginner all his life. - Collingwood [COL42]

Many diverse organizations rely on information systems that depend heavily on distributed legacy data sources. These organizations require quick and efficient visualization and manipulation of data with numerous distinct data structures. These data sources introduce a number of significant problems, especially when they must be combined and displayed to remote users. Nielsen [NIE94]; Tullis [TUL93], and many other researchers have investigated this vital issue with various interface systems. However, empirical studies have not been published that examine remote interfaces to distributed heterogeneous data. Shneiderman [SHN97a] describes remote interface systems to complex distributed data as a topic requiring further investigation. This research determines whether a system that provides a coherent representation of this data is more satisfactory than a traditional system for users to perform representative tasks correctly, confidently, and speedily.

This research contrasts the interaction metaphor of *indirect manipulation* with the metaphor of *direct manipulation* described by Shneiderman [SHN83]; [SHN97b] and with the metaphor of anthropomorphic *interface agents* described by Maes [MAE94a]. Direct manipulation is generally characterized by the following: continuous visibility of the objects of interest; physical actions instead of complex syntax; and rapid, incremental, reversible, immediately visible operations. Currently it is impossible for many complex systems to exhibit all attributes of direct manipulation. This is especially true in a distributed system where a remote interface agent system generates actions to a mediator agent that, in turn, accesses distributed heterogeneous legacy data sources in their original forms. In such cases, the interaction metaphor is called *indirect manipulation*.

The research methodology of using a remote visual interface agent to access and add *coherence* to distributed heterogeneous legacy data sources is applicable to a wide variety

of functional domains as shown by Decker, Sycara, and Williamson [DEC97a], and Alderson and Shah [ALD99]. Coherence denotes similarity between items and a logical, orderly, and aesthetically consistent relationship of parts according to Murphy and Medin [MUR92] and Gröller [GRO92]. The concept of coherence introduced in this research is a characteristic of the interface system that presents components in a logical and consistent manner. This concept affects the efficacy of the information derived from the displayed data, which may be evaluated with a set of usability factors.

A significant problem with legacy data is to enable efficient remote access to information from distributed heterogeneous data sources. There are three components of this problem: (1) accessing the diverse distributed data, (2) providing remote visual interface systems, and (3) adding coherence to the resulting information.

This research addressed these issues by extending an agent mediator system to access the data and developing three new research systems. The first research system provides a remote interface to the agent-mediated data sources and represents a traditional system. The second system relies on usability theory to enhance the traditional methodology. This enhanced interface system is designed to provide a coherent remote interface to the agent-mediated data. A third system was developed to provide empirical support for the theoretical solution by enabling remote usability evaluations.

A specific requirement of this research is to determine whether the Visual Interface To Agent Mediated Information Networks (VITAMIN) enhanced methodology is superior to the traditional Java Indirect Manipulation Interface (JIMI) methodology. The first research system, JIMI, was developed as a control treatment to represent a legacy approach with a traditional query interface. VITAMIN, the second research system, was developed to add coherence to the legacy data with an enhanced interface system.

The JIMI system and the VITAMIN system are implemented as *indirect manipulation* interface systems (IMIS) or simple, non-anthropomorphic interface agents. They are shown Figures 1 and 2, respectively as prototype, stand alone, Java interface systems.

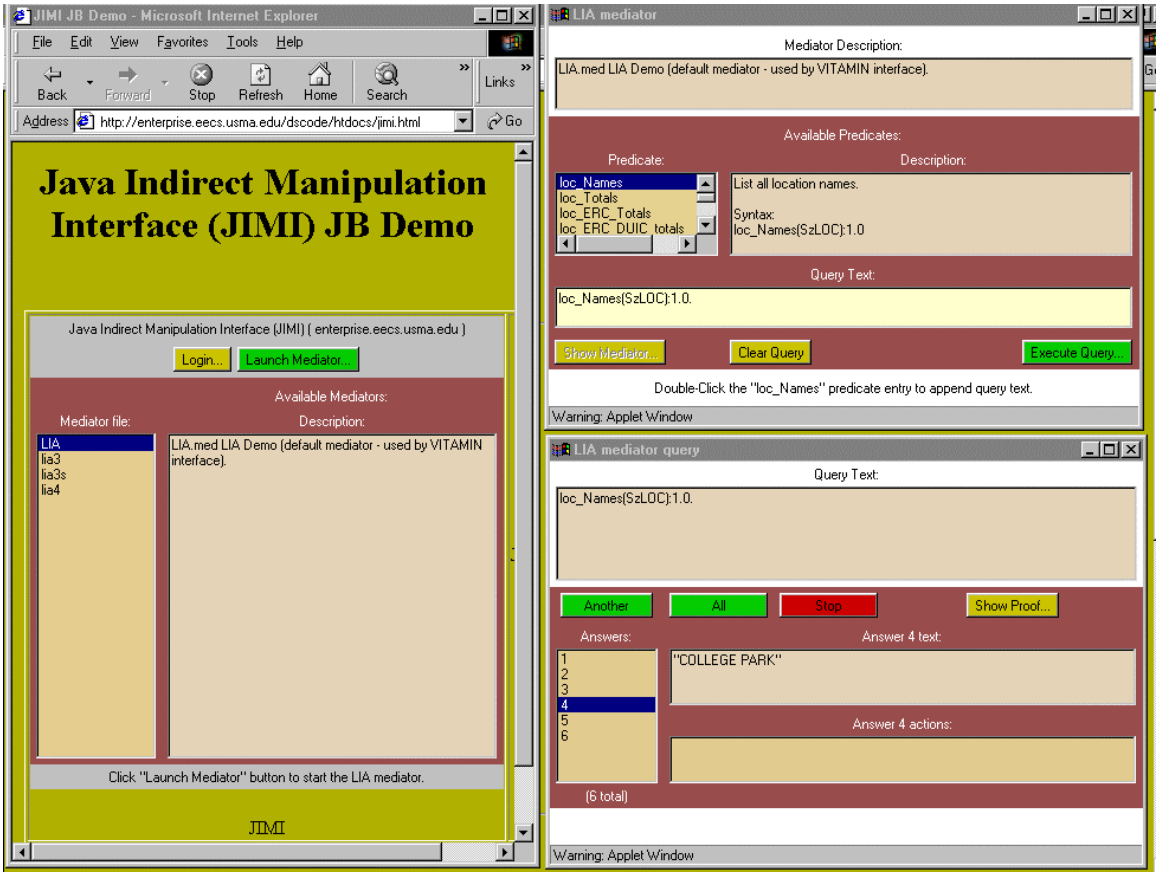


Figure 1 JIMI Screen Capture

The JIMI and VITAMIN systems provide intermediate query predicate actions to the Heterogeneous Reasoning and Mediator System, HERMES, designed by Subrahmanian et al. [SUB97a]. HERMES is an agent mediator system in the IMPACT agent architecture that provides access to legacy data sources for this research. IMPACT is the Interactive Maryland Platform for Agents Collaborating Together described by Arisha et al. [ARI98].

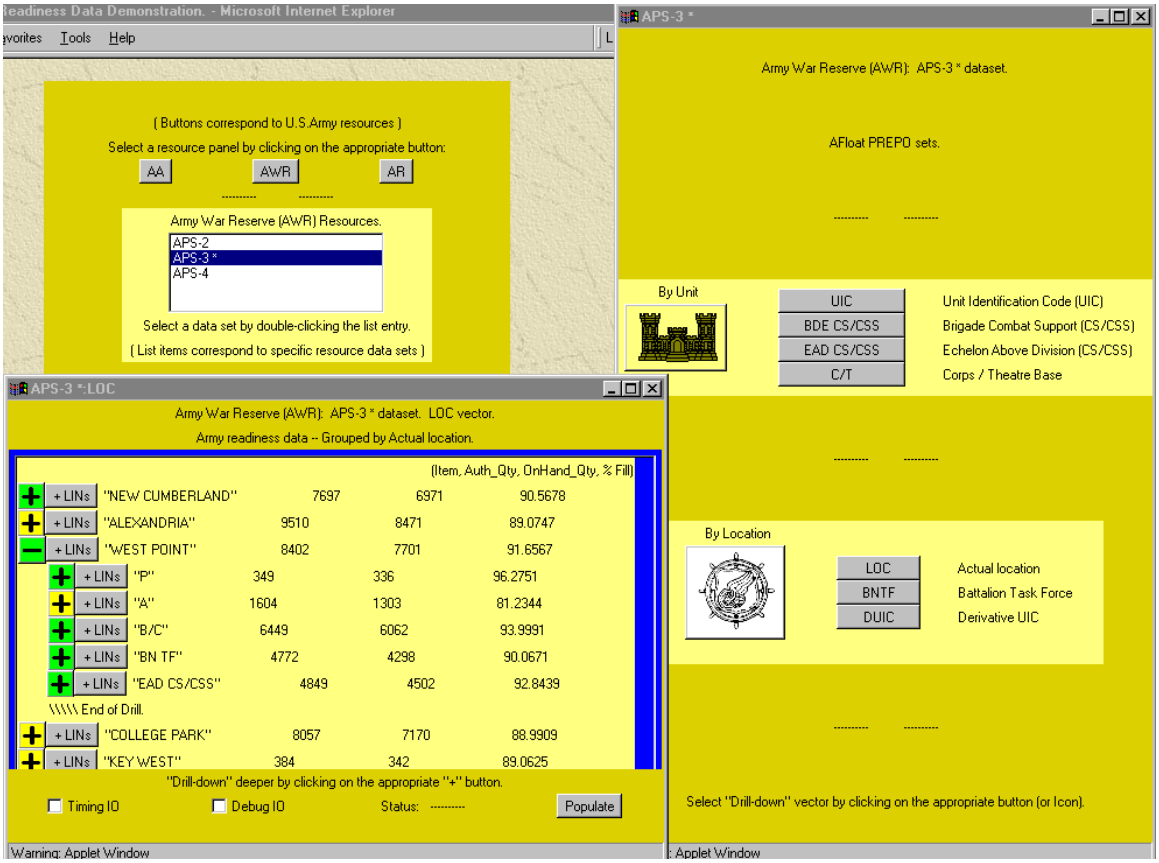


Figure 2 VITAMIN Screen Capture

The remote usability experiment is enabled with the Java Usability Interface Comparison and Evaluation (JUICE) system. JUICE was also developed for this dissertation. JUICE is a unique, remote evaluation system. JUICE uses HERMES to manage the experiments and record the experimental results. The relationship between these systems is shown in Figure 3. HERMES provides a good approach for mediating between diverse data sources and reasoning systems. Additionally, this researcher had access to the HERMES source code to make necessary modifications to interface with complex legacy data sources.

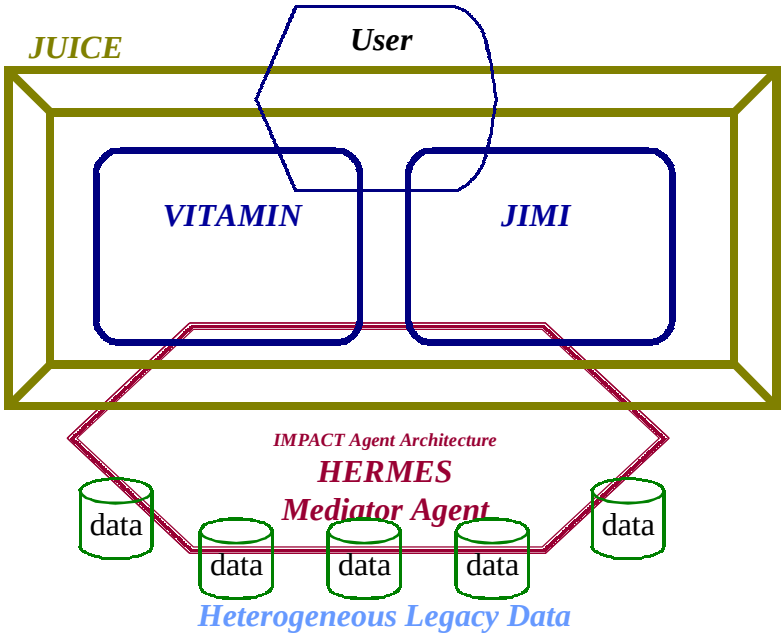


Figure 3 Research Systems Relationships

Remote usability evaluation has been broadly described by Hartson et al. [HAR96b]. This dissertation presents the results of remote usability experiments. These web-based experiments empirically determine whether coherence is enhanced through the application of the research methodology by presenting one interface system and a sequence of representative tasks. Figure 3 depicts JUICE wrapping VITAMIN and JIMI. JUICE presents the IMIS and the set of representative tasks. The subject uses the IMIS to answer the tasks. The IMIS uses HERMES to mediate the heterogeneous data sources and JUICE records the information answers and other information in a data source accessed through HERMES. The JUICE remote evaluation system measures coherence based upon the subject's time to complete each task, the correctness of their answer, and

their subjective confidence in that answer. When all tasks have been completed, the users complete a usability survey to express their satisfaction with the interface system. The main empirical effort of this research is a comparison of the two interface systems developed for this study.

The functional area selected for this research is information operations. This dissertation is specifically concerned with the military logistics field of Army War Reserve (AWR) equipment readiness. This research considers the readiness of the set of AWR equipment that is prepositioned in task-organized units aboard various ships around the globe. An introduction to the complexity of this domain is presented in section 2.5, *Heterogeneous Legacy Data Sources*. This work has broad applicability to myriad remote interface systems that must interface with complex distributed data. In addition to the work presented in this dissertation, the author also applied this methodology to the equally challenging domain of combat simulations as reported by Byrnes, Schafer, and Marin [BYR99].

1.1 Domain Motivation

AWR equipment readiness falls in the realm of logistics planning in a military environment. Logistical planning for the 21st Century Army requires the quick and efficient visualization and manipulation of distributed legacy databases with diverse forms of data structures. Since these data sources are in different locations, on different platforms, and stored in different database or file formats, this becomes an extremely difficult problem. Currently, in a labor- and time-intensive effort, Army War Reserve (AWR) analysts manually combine such sources into a common format on a common platform. According to White [WHI98], expert analysts then issue queries against the aggregated data snapshot from a traditional command line interface.

This traditional approach is compared with the research methodology that aims to provide a more coherent interface to the legacy data. In order to evaluate the interface system differences properly, JIMI differs from the traditional legacy interface system in two ways. First, JIMI uses HERMES to access the data instead of using a single snapshot

homogeneous database. Secondly, JIMI displays example query predicate actions to enable non-expert subjects to participate in the experiments.

1.2 Research Problem

Inspired by the work of Shneiderman and Maes [SHN97c], Lewis [LEW95], and Hartson et al. [HAR96b] on direct manipulation and interface agents, succinct usability metrics, and remote evaluation, respectively, this dissertation research asks the following question:

Can an enhanced indirect manipulation interface system add coherence to agent-mediated legacy data for users performing representative tasks?

An enhanced *indirect manipulation* interface system (IMIS) is one that is superior due to visual cues that allow the data to be presented more coherently. These visual cues include improved data selection instead of data entry, improved visual clarity with familiar color effects, and improved representation with an aggregating tree structure. Adding *coherence* to the data means that the information efficacy is increased such that the usability and usefulness of the system has been improved. Derived from the IMIS, coherence is succinctly measured by (1) the number of correct task answers, (2) the user confidence in the task answer, (3) the time to answer a task, and (4) the user satisfaction with the interface system. The agent-mediated data sources are approached in a particular context that is determined by the specific users, tasks, and environment. Representative tasks based upon interviews with expert domain analysts; they are questions that analysts must commonly answer from the legacy heterogeneous data sources.

1.3 Research Methodology

The initial stage of this research involved identifying and collecting characteristic samples of the heterogeneous data. The next stage necessitated understanding the data and processes in order to engineer data interfaces and mediators for HERMES. Users and experts where interviewed and the JIMI and VITAMIN interface systems were developed. According to Plaice [PLA95] and Tichy [TIC98], remote usability testing and

experimentation is vitally important to the field of computer science. JUICE was developed to enable a remote empirical comparison of the two interface systems. A formative evaluation of JUICE itself was conducted during the pilot study.

VITAMIN and JIMI were developed as stand alone interface systems. For the experimental system, JUICE encapsulates these interface systems by presenting the interface system on the left side of the screen. Representative tasks and survey questions were presented on the right side.

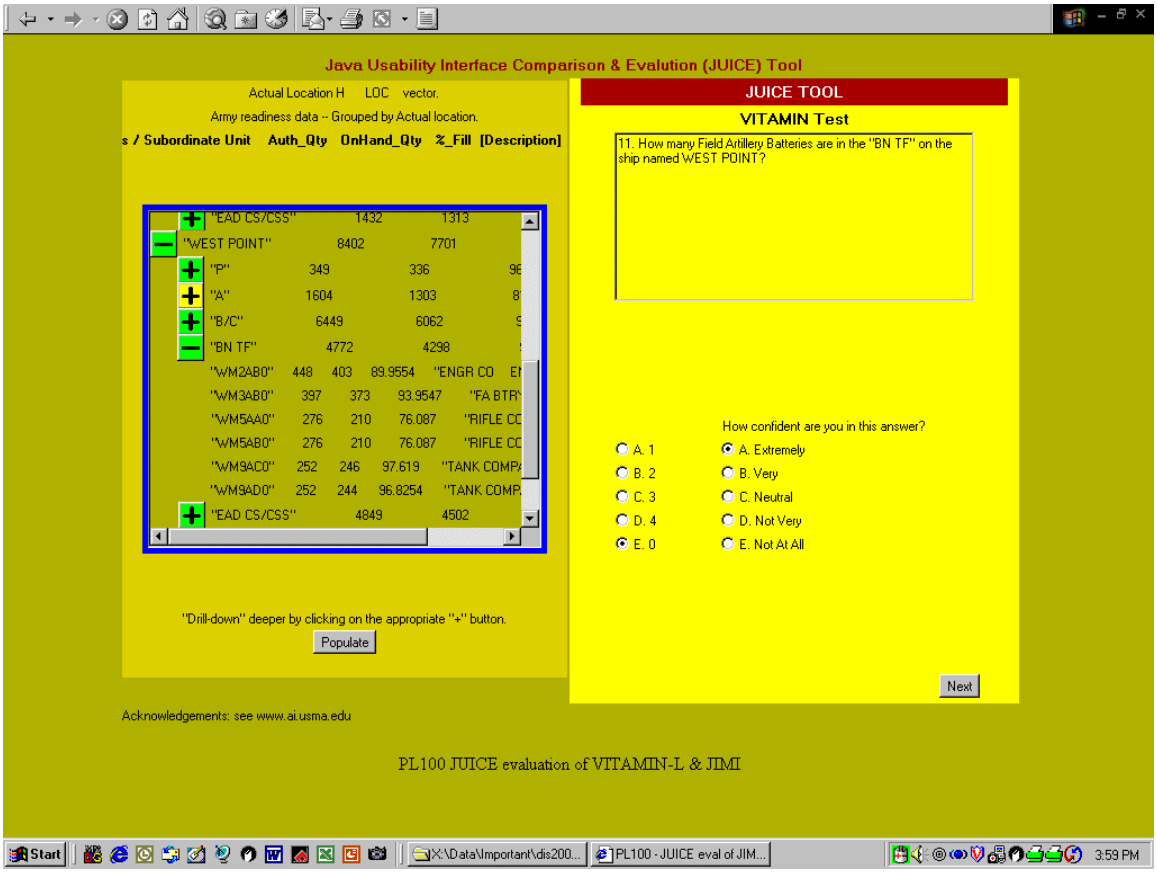


Figure 4 JUICE Task with VITAMIN

During the research experiment, subjects access JUICE from the research web site. Figure 4 illustrates JUICE presenting VITAMIN in the left panel and a representative task in the right panel. JUICE sequentially presents the entire set of twelve tasks on the right side, and either the VITAMIN or JIMI system on the left side.

Each subject was tested with a set of representative tasks on both IMIS treatments, VITAMIN and JIMI. JUICE randomly assigns the treatment order to enable a counterbalanced, within-subject experimental design. The subjects use VITAMIN or JIMI to determine the correct answer for each task. VITAMIN or JIMI then passes the resulting predicate actions to HERMES.

HERMES accesses and mediates the heterogeneous legacy data sources in the IMPACT agent architecture. Then JUICE presents the alternate interface system and another parallel set of tasks. JUICE reports the subject's actions and answers to HERMES. HERMES stores this information on a remote server for later analysis.

Fifty-six subjects participated in the dissertation experiments. Of these, forty-one completed the experiment. Therefore, the within-subject dissertation experiments yielded 82 experimental research cases for analysis. An additional 63 subjects were tested in two phases of pilot experiments. In the dissertation experiments, the representative tasks were divided into three types. This division was based upon the levels of comprehension, application, and analysis required for answering each task according to Bloom's [BLO56] taxonomy.

Statistically significant results were obtained for all four indicators of coherence with VITAMIN providing an improvement over JIMI for each hypothesis: (1) the number of correctly answered tasks; (2) the user confidence in correctly answered tasks; (3) the time to answer tasks correctly; and (4) the user satisfaction with the interface.

Chapter Two, *Literature Review*, assesses previous work that inspires and grounds this research. Interface system usability, usability metrics and remote evaluation, coherence, and indirect manipulation are each reviewed in this next chapter. The functional domain background and software agent theories that enable access to the heterogeneous legacy

data sources are also reviewed. Chapter Two concludes with a summary of the open issues in this research area. Chapter Three, *Methodology*, discusses the research system implementation and the experimental design. Chapter Three incorporates the following sections: research systems and hypotheses, experimental design, and ethical considerations. Chapter Four reports the *Results and Analysis* and Chapter Five, *Conclusions and Future Work*, presents the research lessons and implications for further study.

CHAPTER 2: LITERATURE REVIEW

> You pull one book from the shelf, which carries a hint or a reference that sends you posthaste to another book, and that to successive others. It is incredible, the number of books you hopefully open and disappointedly close, only to take down another with the same result. - Wells [WEL37]

This dissertation relies on usability theories to construct interface and evaluation systems and the research uses an agent-based approach to access the legacy data sources. According to the ACM Task Force on the Core of Computer Science [SCI89], usability and agents are extremely active and important research areas. Additionally, to commemorate the 50th anniversary of the ACM and the computing discipline, over 300 participants gathered at Massachusetts Institute of Technology (MIT) in June 1996 to examine strategic directions in computing research. Computing research was grouped into three broad areas: foundations, systems, and applications and infrastructure. According to Wegner and Doyle [WEG96], the reports of 19 working groups were placed into these three areas. Agents and usability each encompass portions of all three of the broad computing research areas identified above.

A motivating example of the importance of usability research is provided by the experiments conducted by the US Army that focused on enhanced situational awareness. According to Bond [BON98a], the Army conducted several large-scale exercises to test the digitized battlefield in the late 1990's. Digitizing the battlefield means applying information technologies to acquire, exchange, and employ timely digital information throughout the battlespace. This information is tailored to the needs of each decision maker, shooter, and supporter. This wealth of information allows all operators to maintain a clear and accurate vision of the battlespace necessary to support both planning and execution.

The "digitized" forces' superior awareness of both friendly and enemy forces proved invaluable in planning for actions and in reviewing completed actions according to Bond [BON98a]. However, according to the Army's top research and information technology

leaders, Lieutenant Generals Kern [KER98a] and Campbell [CAM98a], in the heat of battle, soldiers and leaders at all levels fell back on traditional manual techniques and did not use their digitized interface systems. The Army's current senior information operations leader, Major General Cuviello [CUV99], goes further and says:

> *The problem is that we don't understand the human factors. We are trying to do digital business in an analog way with digital enablers! Situational awareness is the goal. Not just position and strength, but bullets, fuel, health, and all of combat, combat support, and combat service support. We now have a constant digital feed of information, not just 0600 and 1800 reports. For years we have been saying, "Just keep it coming, I'll know what I need when I see it," - particularly in the intelligence community. Now all battlefield functional areas are feeding data such that a commander may have eight different computers on his desk. We have paralyzed our leaders with data. We have finally moved beyond raw data to information, but we must get to knowledge and improve the human factors.*

These senior leaders voice a frustration ringing from many corners of the information age. The problem is not too little data; the problem is extracting decision knowledge from the vast sources of data available. While most organizations may not have literal life-and-death reliance on interface systems, many rely heavily upon their ability to quickly display answers from heterogeneous networked information sources. There is a substantial need to make vast quantities of information more useable.

This chapter reviews previous work in usability and usability evaluation. A great deal of additional effort was invested to solve the problems of decoding and accessing the heterogeneous data sources. As mentioned in Chapter One, an agent-based approach to accessing this data was selected. A brief appraisal of that research and an explanation of the diverse data sources are presented at the end of this chapter.

Section (2.1) describes the meaning of usability in a computer interface system context. The next three sections review research on (2.2) usability metrics and remote evaluation, (2.3) *coherence*, and (2.4) *indirect manipulation*. Section (2.5) describes the supporting literature related to the functional domain and software agents. The final section (2.6) summarizes and analyzes this past work for pertinent open issues.

2.1 Interface System Usability

> Information networks straddle the world. Nothing remains concealed. But the sheer volume of information dissolves the information. We are unable to take it all in. - Tannen [TAN98]

Helping people to make sense of vast volumes of information is the role of interface systems usability. In the context of user interface systems, the term *usability* was coined in the early 1980s to replace the term *user-friendly*, which, according to Bevan, Kirakowski, and Maissel [BEV91], had become vague and overused. This section reviews attempts to define the abstract concept of usability in terms that are more precise. Standard precise usability factors are summarized in Figure 5.

Effectiveness	Accuracy Completeness
Efficiency	Time Resources Human Resources Financial Resources
Satisfaction	Comfort Acceptability

Figure 5 ISO Usability Factors

The International Standards Organization standard 9241 defines usability as the "**effectiveness**, **efficiency** and **satisfaction** with which a **specified set of users** can achieve a **specified set of tasks** in a **particular environment**," ISO [ISO91] (emphasis added). Once the users' tasks are identified, effectiveness, efficiency, and satisfaction

must be decomposed into measurable attributes. *Effectiveness* is the accuracy and completeness by which specified users can achieve specified goals in particular environments. *Efficiency* is the resources expended in relation to the accuracy and completeness of goals achieved. These resources include time, people, and money. *Satisfaction* is the comfort and acceptability of the work system to its users and other people affected by its use.

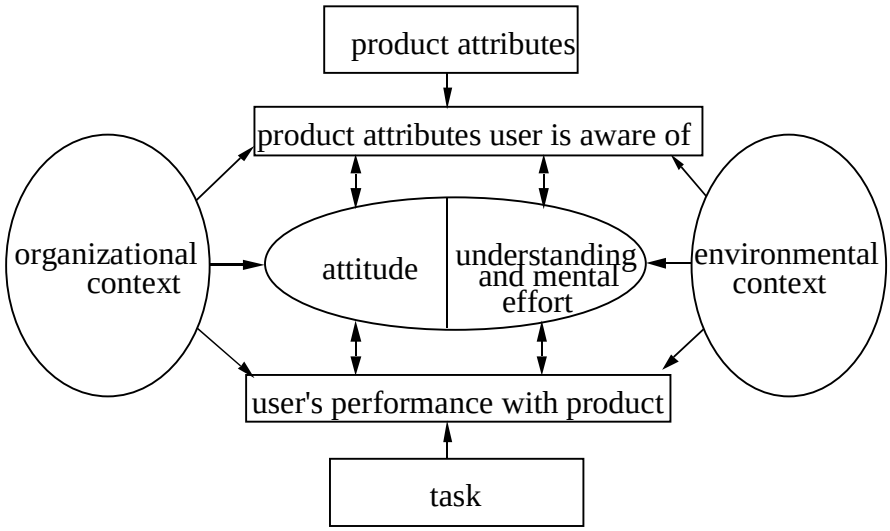


Figure 6 MUSiC Determinants - Bevan, Kirakowski, and Maissel [BEV91]

Another international effort to define usability is the Measuring Usability of Software in Context (MUSiC) project, HFRG [HFR93]. According to MUSiC, usability is defined in terms of ease of use to include ease of learning, acceptability, and actual usage of a specific user for a specific task in a specific context. Ease of use determines whether a product can be used. Acceptability determines whether it will be used and how it will be used. Ease of use in a particular context is determined by product attributes, and is measured by user performance and satisfaction. The context consists of the user, task, and environment. The relationship between these factors is shown in Figure 6.

It is interesting to note the general agreement in the research community that usability cannot be separated from the overall context. This context includes the users, the tasks, and the overall organizational and environmental situation. The primary researchers involved in the International Standards Organization's work on usability offer this description of various aspects and definitions that highlights the *context of the use* across several views of usability:

> *There are still many different approaches to making a product usable, and no accepted definition of the term usability. The definitions which have been used derive from a number of views of what usability is. Three of the views relate to how usability should be measured:*
>
> - *the product-oriented view, that usability can be measured in terms of the ergonomic attributes of the product*
> - *the user-oriented view, that usability can be measured in terms of the mental effort and attitude of the user*
> - *the user performance view, that usability can be measured by examining how the user interacts with the product, with particular emphasis on either ease-of-use: how easy the product is to use, or acceptability: whether the product will be used in the real world.*
>
> *These views are complemented by the contextually-oriented view, that usability of a product is a function of the particular user or class of users being studied, the task they perform, and environment in which they work. Bevan, Kirakowski, and Maissel [BEV91]*

The relationship of these factors and views is displayed in Figure 7.

To this point, discussion has been limited to the definition of usability and its various factors. The primary point is that the somewhat abstract concept of *user-friendliness* or *usability* can be decomposed into more precise components that may be systematically approached. The next step is to address usability evaluation.

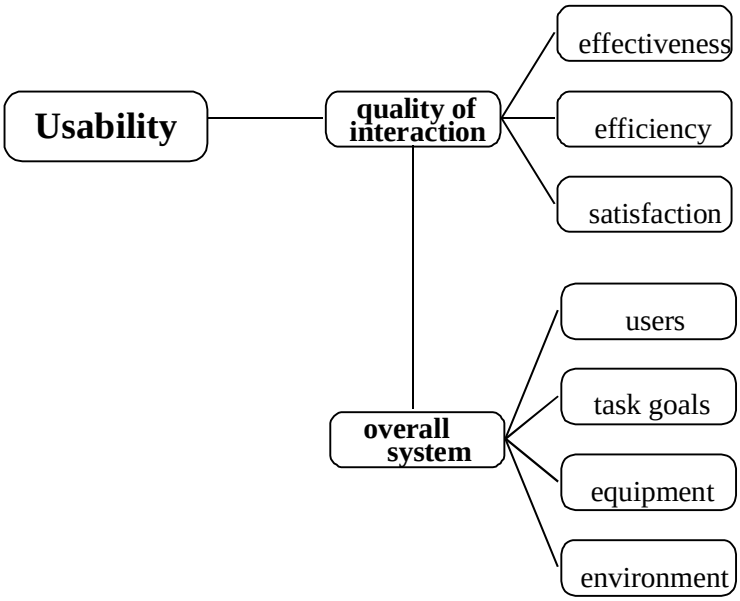


Figure 7 Usability Factor Relationships - Bevan and Macleod [BEV93]

Concentrating on the **quality of interaction** (in Figure 7) for usability evaluation, these standard usability factors may be mapped to measurable attributes in the **overall system** as follows:

- Effectiveness: Number of tasks answered correctly
- Efficiency: Time and number of actions required to complete each task
- Satisfaction: Confidence that the system provided the correct answer and an overall feeling that the system was helpful, was quick, provided quality information, had a quality interface, was easy to learn, was enjoyable, and was useful.

For the remainder of this work, *satisfaction* refers to overall *system satisfaction* and task *confidence* is treated separately. The next section looks at the background of these

attributes and other aspects of measuring usability. Specific evaluation instruments and considerations for remote evaluations are addressed.

2.2 Usability Metrics and Remote Evaluation

> …there is no such thing as a unique scientific vision ... science is a mosaic of partial and conflicting visions… - Dyson [DYS95]

Usability evaluation to determine system quality and user acceptance provides a concrete basis to create and improve truly usable systems according to the Human Factors Research Group in Ireland, HFRG [HFR93]. To address the most glaring usability issues during product development, Virzi [VIR92] found that a very small number of subjects (4-5) were able to account for 80% of disruptive usability problems. Tests of small numbers of subjects on stand-alone applications are well served by a usability laboratory. However, according to Spyridakis [SPY92], many more subjects are required if statistical comparisons of different systems or methodologies are examined. Remote evaluation techniques offer the ability to test a large number of subjects. Additionally, "the network itself and the remote work setting have become intrinsic parts of usage patterns, difficult to reproduce in a laboratory setting" according to Hartson et al. [HAR96b]. Therefore, in many current settings, remote evaluation offers an opportunity for more realistic results than laboratory evaluation.

According to Nielsen [NIE94], the primary usability methods include: heuristic evaluations, performance measures, thinking aloud protocols, observation, questionnaires, interviews, focus groups, logging actual use, and user feedback. Several of these metrics must currently be administered in the relatively controlled environment of a usability laboratory. However, as network bandwidth increases all of these methods may eventually be available for remote administration. Three usability metrics that can now be administered remotely include: *performance measures, logging actual use,* and *questionnaires* according to Hartson et al. [HAR96b]. These metrics are now examined in turn.

Logging actual use, described by Nielsen [NIE94], refers to collecting data based on the users' actions. The computer or evaluation system automatically collects statistics about the detailed use of the system. Logging is a straightforward process that may also be used to capture additional *measures of performance*. In a remote setting, this can be accomplished by an instrumented remote evaluation. Once this performance data has been collected, the next step is to analyze the data with various indicators of usability based on performance. Hartson et al. [HAR96b] describes instrumented remote evaluation as using metering code embedded into the application. Often, however, researchers may not be able to embed additional code into applications, suggesting the need for another approach.

In a survey of HCI literature, Rengger [REN91] identified four classes of these *performance measures*: goal achievement (accuracy and *effectiveness*), work rate (productivity and *efficiency*), knowledge acquisition (learnability and learning rate), and operability (error rate and function usage). Considering *quality of interaction* from above, performance measures of log data can be used to measure *effectiveness* and *efficiency*. *Satisfaction* and *confidence*, on the other hand, are more difficult to measure with log data. One solution is to use logging as either a stand-alone technique to unobtrusively gather usage data from subjects or as part of more formal evaluations that use task *questionnaires* and surveys.

Questionnaires and surveys to measure user satisfaction can be very complex and sophisticated as shown in Nielsen [NIE94]. Two such prominent usability survey metrics are the Technology Acceptance Model (TAM) described by Davis, Bagozzi, and Warshaw [DAV89], and the Goals, Operators, Methods, and Selection Rules (GOMS) method by Card, Moran, and Newell [CAR83]. Several additional usability questionnaires also aim to achieve high reliability and validity across a wide variety of interfaces. Two such *de facto* standard instruments are QUIS by Chin, Diehl, and Norman [CHI87]; [CHI88], and SUMI by Kirakowski and Corbett [KIR93]; Porteous, Kirakowski, and Corbett [POR93].

TAMS is an attitude questionnaire to predict end-user acceptance of applications. The model includes **perceived usefulness**, *"the degree to which an individual believes that using a particular system would enhance his or her job performance"* and **perceived ease of use**, *"the degree to which an individual believes that using a particular system would be free of physical and mental effort"*.

GOMS is an analytical method that attempts to predict usability by describing user tasks. The basic GOMS method involves listing the task *goals* and subgoals, the *operators* available to the users, the *methods* users construct from the operators, and the *selection rules* for the deciding upon the next goal or method. According to Nielsen [NIE94], and Sutcliffe et al. [SUT91], the most important weakness of the GOMS model is that it is primarily limited to expert users. Many modified GOMS models have shown benefits to analyze usability including Mackinlay, Rao, and Card [MAC95] and John and Kieras [JOH96b].

- Screen factors
- Terminology and system feedback
- Learning factors
- System capabilities
- Technical manuals
- On-line tutorials
- Multimedia
- Voice recognition
- Virtual environments
- Internet access
- Software installation

Figure 8 QUIS component groups - LAP [LAP99]

QUIS requires users to rate 27 attributes on a 10-point scale. QUIS includes a demographic questionnaire and a measure of overall system satisfaction along six scales, and eleven specific interface factors as shown in Figure 8. Each QUIS area measures the users' overall satisfaction with that component of the interface, as well as the factors that make up that component according to HCIL [HCI99].

SUMI users rate five groups of 50 attributes on a 3-point scale (agree, undecided, disagree). These groups are described in Figure 9. Similar to these groups, Nielsen [NIE94], identifies the following five usability attributes: Learnability, efficiency, memorability, errors, and satisfaction. SUMI's efficiency and learnability correspond to Nielsen's like-named attributes. *Affect* corresponds to satisfaction. Nielsen's *errors* attribute measures avoidance and recovery from user mistakes and is subsumed by SUMI's notion of control. Memorability is related to helpfulness insofar as the casual user should not have to relearn the system from scratch.

Attribute	SUMI Meaning
Affect	Subject's emotional feeling toward the software
Efficiency	Sense of degree to which the software enables the task to be completed in a timely, effective, and economical manner
Learnability	Feeling that it is straightforward to become familiar with the software
Helpfulness	Perception that the software communicates in a helpful way to assist in the resolution of difficulties
Control	Feeling that the software responds to the user inputs in a consistent way and that its workings can be internalized

Figure 9 SUMI Measurement Scales - HFRG [HFR99]

The most important contribution of the TAMS, GOMS, QUIS, and SUMI methods to this research is the identification and clarification of measurable usability factors. Despite the proven breadth, reliability, and validity of these comprehensive instruments, several researchers including Nielsen [NIE94], and Rubin [RUB94], recommend shorter questionnaires. Lewis [LEW92] found in a factor analysis of an 18-question satisfaction questionnaire that 87% of the total variance in responses was due to only three factors: system usefulness (correct and quick), information quality (correct and confident), and interface quality (satisfactory).

These factors, and the small number of important factors identified by other researchers, point to the use of much simpler, tailored instruments as shown by Brooke [BRO96]. Additionally, Nielsen [NIE94] describes *discount usability engineering* which argues for "quick-and-dirty," non-statistical usability analyses. Simpler instruments are also much more manageable in a remote evaluation system where the researcher may have less influence over the experimental environment than they would have in a usability lab.

The discussion thus far has shown that *effectiveness* and *efficiency* may be measured with performance measures of log data, and *confidence* and *satisfaction* may be measured with short survey instruments. Additionally, the point has been made that remote usability evaluations may be more representative for certain networked applications than laboratory evaluations. The next question is: How may these evaluations be conducted?

Web-based remote experiments were conducted by Nebesh [NEB97], and Samadi [SAM97]. They demonstrated that experiments could be successfully accomplished using the web. These experiments were not usability evaluations, however, they did compare various systems and record questionnaire and timing information. Perlman [PER98] developed a customizable Web-based perl CGI script to administer and collect data according to standard user interface evaluation questionnaires such as QUIS and SUMI. Perlman's script can be used to evaluate web pages. Subjects' responses are emailed to the researcher in a simple format. The script has no option to record task response times or present various systems for comparison. No systems to enable these evaluation requirements were discovered during this research.

This section recognized the opportunity for a unique, empirical, remote usability evaluation system. Additionally, four important usability factors: effectiveness, efficiency, confidence, and satisfaction have been identified. The evaluation system must select an interface system to be assessed, present representative tasks to users, record their answers (*effectiveness*), time to answer (*efficiency*), and *confidence*. For a counter-balanced design (CBD), the evaluation system must then select another interface system from the available alternative interface systems. The evaluation system is again required to present tasks and record responses for the alternate interface system. Then the evaluation systems must administer a survey instrument to compare the interface systems and assess the overall *satisfaction*. No example of a remote evaluation system to provide a controlled, task-oriented, statistical comparison of two interfaces was found in the literature. Thus, an opportunity to specify an innovative remote usability system has been identified.

The next section defines a concept for describing the improved information efficacy provided by enhanced interface systems. This concept, *coherence*, is also introduced here as a term to described a concise usability metric for statistical analysis.

2.3 Coherence

> Perhaps believing in good design is like believing in God, it makes you an optimist. - Conran [CON89]

Coherence is defined here as a characteristic of interface systems that present interaction components in a logical and consistent manner. The term *coherence* is used in this research as both a theoretical framework and an implementation metric. Coherence affects the efficacy of the information that a user derives from displayed data. It is important to note that enhancements to interface systems frequently improve coherence; therefore, the concept of coherence is a characteristic of the interface systems and not a characteristic of the data. An indirect manipulation interface system (IMIS) that has improved the information efficacy in a specific, well-defined context enhances

coherence. Context refers to a specific set of users and tasks in a specific environment. This improvement may be evaluated according to important usability factors identified in the previous section.

These factors include the performance of specific tasks *correctly, confidently, speedily, and satisfactorily* (C^2S^2). Coherence is also introduced in this research as a term to describe a concise usability metric for statistical analysis. This broad characterization of coherence is established to denote these four usability factors without requiring a precise set of interface devices or metaphors. Coherence is particularly appropriate for remote evaluations as a succinct metric because each of these factors may be considered by remote usability evaluation systems.

An interface system that is enhanced to facilitate these factors is more coherent than one that is not. Perlman [PER94b] describes such interface system enhancements. The concept of coherence allows straightforward comparison of a traditional interface system with an enhanced interface system. This concept will be further examined in this section. The remainder of this section is organized as follows: interface systems design and the meaning of coherence in other contexts are considered, interface systems enhancements are reviewed, and finally coherence is portrayed in measurable terms.

As with usability metrics, discussed in the previous sections, much is known about good interface system design. This knowledge ranges from the "magical seven plus or minus two" limits of short-term memory to a complete multimedia taxonomy. The limits of short-term memory originally described by Miller [MIL56] have been applied many times to interface designs. Tests based upon the "magical seven" include breadth and depth of menu arrangements, web page links, and information retrieval from an electronic encyclopedia according to Larson and Czerwinski [LAR98]. At the other end of the range, Heller and Martin [HEL95] present a taxonomy that illustrates a comprehensive approach to research and development of multimedia applications and interface systems. This taxonomy enables a common language to describe interface designs.

The discussions in the previous two sections listed many aspects of usability and usability metrics. Like taxonomies and validated usability instruments, many of these aspects are very broad and comprehensive. On the other hand, empirical evidence, also discussed above, suggests that a very simple set of measurements can account for the vast majority of issues. In this vein, a simple set of usability metrics termed *coherence* is defined here as a concise usability metric for statistical analysis.

Coherence denotes consistent similarity between items. In general, coherence is:

> *The quality or state of cohering, especially a logical, orderly, and aesthetically consistent relationship of parts* - [AHD92].

From a psychological perspective, Murphy and Medin [MUR92] argue "theories about the world highlight properties that make a concept coherent." In other words, coherence derives from a certain worldview. Somewhat closer to the interest of this research, Jain, Manuel, and Singh [JAI99], explain that although data consistency is not fully achievable in an open environment with autonomous agents, "a coherent state in the ongoing interactions of the participating components" is absolutely required. The implications are that the software agents must not have unrestrained autonomy in an open environment because they cannot control the actions of other autonomous components. Kermarrec, Steen, and Tanenbaum [KER97] address similar issues with regard to consistent replication of web components. In another usage from the literature, Johnston and Agarwal [JOH95] said that an absolute requirement for remote evaluation was "interface coherence."

In computer graphics, coherence describes the extent to which graphical items or entities are locally constant, according to Gröller [GRO92]. Gröller and Purgathofer [GRO95] further state that coherence techniques may be exploited to increase efficiency. This description from the field of computer graphics is much closer to the interest of this research. Leveraging this work, enhancements to the coherence of interface systems is now examined.

The literature includes research on many enhancements that may be made to interface systems to improve their design and usability. For instance, recent work by Sweeney-Jackson [SWE98] analyzes the legibility of web interfaces based upon text and color. Among many enhancements that may be made, two obvious examples include interaction style and color combinations. Research in these two areas will now be reviewed.

The history of users' interactions with computers has evolved from a relationship where users interacted only with the computer (and not directly with the task) to the present situation where users often interact more directly with tasks. According to Norman [NOR88], this more direct task interaction reduces the mental distance. This distance has reduced as interaction styles have progressed from batch processing, to line entry, to full screen operations. Tullis [TUL85] described research on menu-based systems. Full screen operations have progressed from character-only and menu based systems to the current systems built around the Windows, Icons, Menus, Pointing Devices (WIMP) style. van Dam [VAN97a] traces this development and argues for a next generation of Post-WIMP three-dimensional interaction styles. Currently, familiar practical interaction metaphors that use WIMP interfaces include direct manipulation and interface agents. These metaphors are contrasted in the next section.

WIMP interfaces often include the familiar tree or drill-down-table metaphor as described in Goldstein and Roth [GOL94]. A common example of this metaphor is the file manager or explorer available in many graphical operating system interfaces. This summarizing or aggregating tree typically presents an overview of complex data that is represented by a tree of plusses and minuses that expand and contract, to more or less detailed information. According to Erickson [ERI90], a metaphor that adds structure such as a summarizing tree is enhanced compared to one that requires text entry.

Kristof and Satran [KRI95] and Tullis [TUL93] offer many design suggestions, including the use of color to enhance the user interaction. Colors that are symbolic or familiar for some domain meaning can be especially helpful. For instance, a red button could mean stop and a green button could mean go. A poor design may have these reversed. Tufte [TUF90] offers four principles of color for good visual effects. (1) Strong bright colors

are best used sparingly or against dull backgrounds to bring attention to the strong color. (2) Mix light and dark colors to provide contrast and emphasis. (3) Backgrounds should be muted to allow smaller bright areas to stand out. (4) Large areas with separate colors should intermingle with each other to impress the mood upon the viewer and prevent a visual clash. According to Jordan [JOR98], an interface design that improves visual clarity with visual cues such as a good color scheme is enhanced compared to one that does not.

For a given task, an interface system with meaningful color and a summarizing tree structure is said to be enhanced over a more traditional system that does not leverage color and that requires text entry. To quantify the enhancement, the concept of coherence must be matched with a succinct set of usability attributes. Recall that, as defined here, coherence comprises the ability to perform representative tasks: *Correctly*, *Confidently*, *Speedily*, and *Satisfactorily* (C^2S^2). The relationship of these attributes to the usability metrics discussed in the previous section is shown in Figure 10.

Tasks Performed	Usability Metric
• Correctly	• Effectiveness
• Confidently	• Confidence
• Speedily	• Efficiency
• Satisfactorily	• Satisfaction

Figure 10 Coherence Metrics

Coherence denotes these four usability factors and does not necessitate a particular set of interface components or styles. For instance, Tullis and Kodimer [TUL95] showed that a command line data-entry technique was much faster than direct manipulation techniques in an empirical evaluation of procedures for sorting data in tables. Specifically, drag-and-drop and menu-selection interface systems were not among the best performing techniques. Therefore, interface systems using these enhanced approaches did not add coherence for these sorting tasks. In this context, command line interface systems were more coherent.

This robust and theoretically derived concept of coherence, identified here, allows the statistical comparison of interface systems consistent with these succinct factors. Accordingly, an interface system is more coherent if empirical analysis reveals statistically significant improvements in these factors. Empirical analysis of specific tasks in specific environments is needed because, according to Jordan [JOR98], grounded usability studies continue to reveal unexpected results that are at odds with theoretical designs. An example of a text-entry system outperforming menu-selection and drag-and-drop interfaces was described above. Additionally, Nielsen [NIE94] describes findings that users often indicate subjective preferences that disagree with objective performance measures.

This section introduced the concept of coherence in relation to interface systems. A useful remote evaluation system may measure coherence based upon (1) the *correctness* of their answer, (2) the subject's subjective *confidence* in that answer, and (3) the time to complete each task (*speed*). When all tasks have been completed, a usability survey of the subjects may record (4) the overall *satisfaction* with the interface system. The discussion in this section included researchers that described coherence in other contexts, metrics to measure coherence, and the reason why the concept of coherence adds to this discussion. Additionally, the literature review revealed no grounded, statistical, remote evaluations of interface systems to agent-mediated distributed heterogeneous legacy data sources.

In the next section, practical interaction metaphors including *direct manipulation* and *interface agents* are examined. The competition of these two metaphors and a theoretical gap between them results in the introduction of a compromise metaphor called *indirect manipulation*.

2.4 Indirect Manipulation

> If we try to squeeze science into a single viewpoint ... we are like Procrustes chopping off the feet of his guests when they do not fit on the bed. - Dyson [DYS95]

Today, the dominant interaction metaphor is called *direct manipulation*. A newer, competing metaphor is called *interface agents*. Much has been made in the literature of the conflict between *direct manipulation* and *interface agents*. Lively debates held at IUI97 and CHI97 between Ben Shneiderman, an advocate of direct manipulation, and Patty Maes, an advocate of interface agents, attracted tremendous attention as reported by Shneiderman and Maes [SHN97c]. Strong and less-than-friendly discussions have been recorded. Shneiderman argues that the user must be able to be responsible for, understand, and control predictable displays. Maes argues that the goal of interface agents is to assist users by filtering information and delegating tasks. This section reviews the history and claims of the competing metaphors and searches for a compromise and common middle ground. First, current issues that motivate this search for middle ground are reviewed. These issues are related to the proliferation of enormous distributed heterogeneous data sources.

According to Codd [COD70] and Stoffel, Taylor, and Hendler [STO97b], retrieving information from huge data repositories in computerized environments has long been an important issue in computer science. Introduction of networked information systems and continuous changes in the amount, type, and format of data make the problems even more challenging as shown in Goh, Madnick, and Siegel [GOH94], and Wiederhold [WIE92].

Obtaining data from distributed data sources via the World Wide Web has become second nature to many users according to Lawrence and Giles [LAW98] and Young and Maracaccio [YOU94]. Yet, effective interaction with distributed heterogeneous data sources means providing the user with access and location transparency as discussed in Hassall et al. [HAS99]. That is, the user should not be concerned with the physical format or location of the data. One major challenge in obtaining transparency is the interactive integration and presentation of heterogeneous, distributed data with a user-friendly, web-based remote interface system according to Caftori, Borenstein, and Hoelscher [CAF98], and Yee [YEE91]. According to Huhns and Singh [HUH98b], one promising approach to achieve this goal is through the combination of interface and mediator agents.

It is a common observation in the human-computer interaction field that the efficient use of a system is also highly related to the user interface of that system. This directly affects the performance of the system itself as described by Davis [DAV93]. These observations lead many researchers to work on the current problems in information retrieval from the user interface side rather than only the processing power side of the system.

The main points thus far highlight the need for improved access to distributed legacy data and the importance of enhancing interface systems to that end. Concentrating on the user interface systems, the dominant interaction metaphors are now reviewed. These metaphors, *direct manipulation* and *interface agents* are compared and synthesized into a new style called *indirect manipulation*.

Ben Shneiderman at the University of Maryland described *direct manipulation* in 1983. According to Shneiderman [SHN83]; [SHN97b], *direct manipulation* is characterized by pointing and selection of objects ("clicking on an icon") to issue actions that are rapid, incremental, and reversible with 100-millisecond updates for all actions. *Direct manipulation* was originally designed for a static, structured, closed, and relatively small information world.

In *direct manipulation*, the WIMP interface portrays a visualization of the objects and actions of interest such that when a user directs an action on the object in the interface, a corresponding action is effected upon the real object. For instance, if a user clicks on a folder icon, the actual directory is opened, the source is listed, and the contents are displayed in an interface window. If the user takes no action, then objects are not changed. *Direct manipulation* has been well researched according to Myers et al. [MYE96a] and Ziegler [ZIE96]. Specifically, Labovitz [LAB67]; Lim, Benbasat, and Todd [LIM96], and Margono and Shneiderman [MAR87] have each conducted empirical comparisons of various direct manipulation techniques. The main characteristics of direct manipulation are summarized in Figure 11.

- Enables visibility of the object and actions of interest
- Allows Rapid, Reversible, Incremental actions
- Replaces complex commands
- Point and select the object of interest

Figure 11 Direct Manipulation - Shneiderman [SHN83]

Although *direct manipulation* is the dominant interaction metaphor in use today, many improvements have been made in the area of computing since its introduction. These technological developments have not significantly changed the way people interact with computers. *Direct manipulation* requires the user to initiate all tasks explicitly and to

monitor all events. According to Maes [MAE94a], this metaphor must change if untrained users are to use of current and future computers and networks more effectively.

Research from the field of Artificial Intelligence, particularly *autonomous agents* techniques, can be used to implement a complementary style of interaction. This complementary metaphor has been called *indirect management* by Dale [DAL97], Gibbins [GIB99], and Kay [KAY90]. The primary research advocate for this metaphor, Patty Maes of MIT, refers to this metaphor as *interface agents*.

Interface agents were designed for a dynamic, unstructured, open, and vast information world according to Maes [MAE94a]. In contrast to direct manipulation where nothing happens until the user clicks on something, personalized *interface agents* may be acting on behalf of the user at all times, like an assistant. This proactivity is possible because the user delegates certain actions to agents that know the user's interests, habits, and preferences. The agents can either act on behalf of the user or make suggestions to her. An example of an anthropomorphic or human-like agent is the animated paperclip that appears as a helper assistant in Microsoft Office products. The main characteristics of interface agents are listed in Figure 12.

- Open, dynamic, unstructured data
- Anthropomorphic
- Know user interests and preferences
- Suggest or act on behalf of user
- Actions without user initiation

Figure 12 Interface Agents - Maes [MAE94a]

The main criticisms of *interface agents* by advocates of *direct manipulation* are that the user must feel in control of a predictable system in order to take responsibility for their actions. Another criticism is that the anthropomorphic representations of agents mislead developers and users by interfering with predictability, control, and responsibility. On the other hand, advocates o*f interface agents* argue that personalized and proactive software is needed to confront the enormous complexity of today's computing environment. According to Jordan [JOR98], interface usability in general has been well researched in homogeneous environments. However, one of the most important problems in this dynamic and complex data world is the ability to generate efficient queries to distributed heterogeneous data sources. Interface agents are designed for this, but direct manipulation researchers have made great progress in this area as well.

Recent user interface research shows that new methods exist for efficient querying on many different environments. For example, the Butterfly and Harvest systems show that efficient querying is possible on the World Wide Web as described respectively by Shneiderman and Maes [SHN97c], and Mackinlay, Rao, and Card [MAC95]. These

systems also attack the problem from the information processing and system utilization sides. Dynamic queries introduced by Veerasamy [VEE95] use a *direct manipulation* approach. Dynamic queries have a visual representation of query components and results. They give continuous feedback to the user for guidance in query formulation.

Dynamic querying is less applicable to huge networked information repositories because of high system-resource requirements, notably bandwidth. A solution to this problem was brought by Shneiderman [SHN93], which uses data aggregation in tandem with dynamic queries. Another solution to the query problem is division into smaller problems such as the design of query previews and overviews as described in Goldstein and Roth [GOL94]. These solutions offer a general overview of the database to the user before the detailed information is requested.

These recent applications of direct manipulation to more dynamic and distributed environments move in the direction of applications for which interface agents once stood alone. The intersection of these two metaphors is defined by this research as *indirect manipulation.* Indirect manipulation uses these results, but does not require the immediate response inherent in *direct manipulation* settings because it is simply not possible in many complex applications. Shneiderman [SHN97a] refers to a similar relaxation of direct manipulation requirements as "remote direct manipulation." Additionally, *indirect manipulation* requires none of the delegation and anthropomorphism of *interface agents*. The features of the indirect manipulation interaction metaphor are summarized in Figure 13.

- Incomplete visibility of objects and actions
- Distributed Data
- Delayed actions
- Remote Devices and Actions
- Delayed or Incomplete feedback

Figure 13 Indirect Manipulation

The term *indirect manipulation* has also been used in the world of Virtual Reality according to Mine [MIN96]. In that context, indirect manipulation means that actions taken on an object are taken indirectly. For instance, if a person wants to hit a table in a virtual environment, she may not be able to do so with her virtual hand, however, she may be able to do so if she picks up a virtual stick. This is a very similar notion to the concept of indirect manipulation described here. The indirect manipulation interface cannot access a data object directly, however, it can generate an action to a mediator agent system that can then access and return a representation of the object. In this way, indirect manipulation interfaces can enable interoperability with distributed heterogeneous data sources.

Like interface agents and recent enhancements in direct manipulation systems, indirect manipulation interface systems (IMIS) are designed for dynamic, unstructured, open, and vast information sources. Distributed data, remote devices and actions, and delayed or incomplete feedback compel indirect manipulation. Direct manipulation provides a tremendous improvement over the previous command-line and menu-based interaction

metaphors. However, the static environment that spawned direct manipulation has changed. Anthropomorphic interface agent interaction metaphors partially address this new distributed and dynamic environment. Between the interaction metaphors of direct manipulation and interface agents, indirect manipulation stakes out the theoretical middle ground.

The next section, *Heterogeneous Legacy Data Sources*, describes the functional data sources and the agent systems that enable the indirect manipulation interface systems (IMIS) to access the data. First, the logistics data is placed within a larger domain and the Army unit organization is illustrated. Then, the software agents, which enable the research data mediation and interoperability, are described.

2.5 Heterogeneous Legacy Data Sources

The functional domain example of military logistics data and processes fall into a larger computing and military domain called Information Operations (IO). A brief explanation of (IO) serves to illustrate the particular importance of usability in this complex environment.

Joint Publication 3-13, Joint Doctrine for Information Operations, the United States Joint Chiefs of Staff defines IO as "actions taken to affect adversary information and information systems while defending one's own information and information systems" JCS [JCS98b]. IO occurs at all times. In time of crisis and conflict, a subset of IO, called "Information Warfare" may be in effect.

One of the key tenets of IO is enhanced situational awareness (SA). SA uses information technology to arrange information about the allied situation and that of the enemy. Unfortunately, current SA systems often rely on poor human-computer interface systems. These interface systems have resulted in lost opportunities and lack of confidence in the systems, thus negating the improved SA.

The US Army Field Manual (FM) for Information Operations, FM 100-6, has this to say about the Army's vision for the Information Age:

> *The Army is embracing a new era characterized by the accelerating growth of information, information sources, and information dissemination capabilities supported by information technology. This new era, the so-called* Information Age, *offers unique opportunities as well as some formidable challenges. New technology will enhance the Army's ability to achieve situational dominance on land, where the decisive element of victory for our nation has always been critical. At the same time, it will enable adversaries to employ many of these same capabilities. This new technology also allows the Army to transform itself.* HQDA [HQD96c]

The specific data sources used by this research are equipment readiness data for Army War Reserve equipment prepositioned aboard ships as described by HQDA [HQD96e]. In order to understand the systemic relationships of the research data, an overview of Army unit organization and *task organization* is presented next.

2.5.1 Army Unit Organization and Data

> Do you know what a soldier is, young man? He's the chap who makes it possible for civilised folk to despise war. - Massie [MAS89]

One of the aspects of this domain that makes it particularly challenging is the concept of *task organization*. The United States Army has an elaborate, bureaucratic hierarchy of commands and units. These units, including descriptions of their training preparedness and equipment readiness for war is the basis of most Army reporting. However, for specific missions, these units are rearranged to provide enhanced capabilities. For instance, an artillery battery may be assigned to an infantry battalion for a particular battle. This is called task organization. Equipment readiness is typically reported through the normally assigned hierarchy. This is true even when a unit is task organized. However, in the case of equipment that is permanently task organized, in prepositioned war reserve configurations, the readiness must be reported by both normal authorized unit and task organized unit. This must be done to present a more accurate representation of the true state of equipment readiness. This requirement greatly complicates the number and type of data sources that must be accessed. The following paragraphs explain the hierarchy and reasons for this complexity.

Army units are typically organized for a mission under a regional theater Commander-in-Chief (CINC) from the top down as follows: numbered armies, corps, divisions,

brigades, battalions, companies, platoons, squads, and teams or sections HQDA [HQD94]. For instance, the Third Army was deployed to Southwest Asia for Operation Desert Storm under CINC US Central Command. Subordinate to Third Army were V Corps and XVIII Airborne Corps (ABC). XVIII ABC commanded 82d Airborne Division, 24th Mechanized Infantry Division, 101st Airborne Division (Air Assault), and several other units. Corps typically have two to five divisions plus Separate Maneuver Brigades and Combat Support (CS) and Combat Service Support (CSS) Brigades and additional units HQDA [HQD96d]. This organizational hierarchy is depicted in Figure 14.

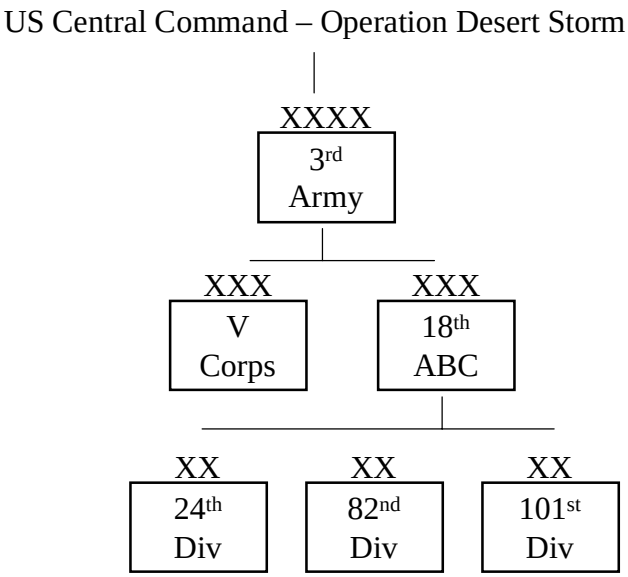


Figure 14 Army Desert Storm Organization

US Army Divisions typically have three combat maneuver (infantry or armor) brigades. Each brigade typically has three battalions, as shown in Figure 15 below, according to Army doctrine, HQDA [HQD96a]; [HQD96b]. Battalions contain several companies. For instance, an armor battalion may have three authorized armor companies and an infantry battalion may have three infantry companies. The companies are represented in

the Figure as A, B, and C Co. For this example, each company has three platoons: 1st, 2nd, and 3rd.

Infantry companies are assigned to infantry battalions and tank companies are assigned to tank battalions. Similarly, tank battalions are assigned to tank brigades and so on. However, mission tasks may dictate that a division commander places a tank battalion into an infantry brigade. This is called *task organization*.

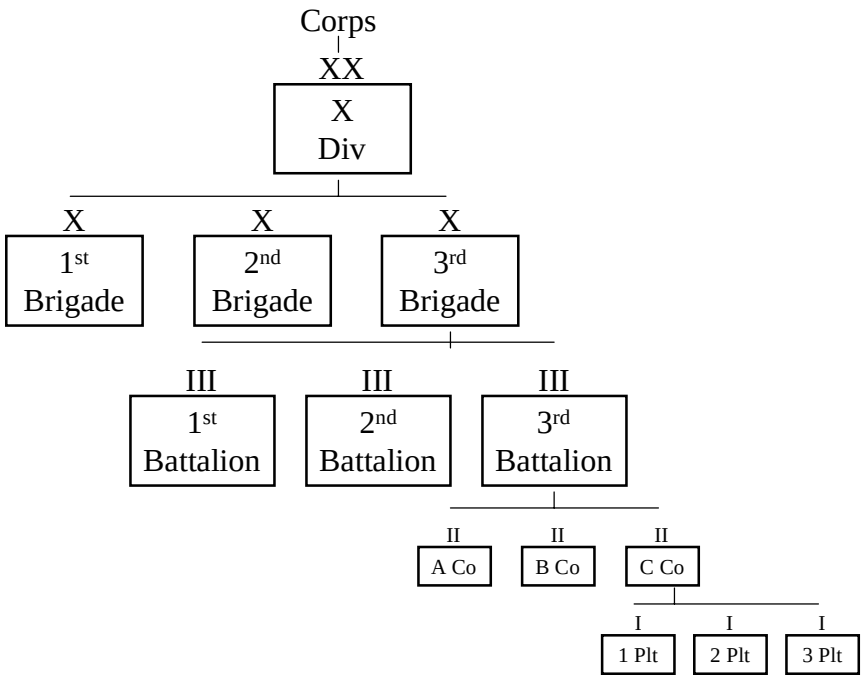


Figure 15 Army Division Organization

In a crisis situation, the Army can generally move many soldiers to any part of the globe very rapidly. The same cannot be said for their equipment. For example, enough soldiers for fifty tanks may fly on a single plane, but that same plane may be able to carry only a single tank. To address this issue, the Army War Reserve (AWR) Command prepositions entire units' worth of equipment at various locations around the world.

These are called Army Prepositioned Stocks (APS). Although the exact configuration of these equipment sets is classified secret, some portion of it is afloat aboard ships. Land-

based sets of prepositioned stocks are referenced by the specific geographical location at which they are stored. The afloat sets of prepositioned stocks are referenced by the name of the ship on which they are stored. The afloat set is called APS-3 for "Army Prepositioned Stocks – Set 3."

For the purpose of this research, it is assumed that the all of the afloat sets together correspond to a separate armor (tank) brigade's worth of equipment. Figure 16 depicts a separate armor brigade. As illustrated above, a brigade is normally subordinate to a division and the division controls the required support units.

A separate brigade is a combat unit of a few thousand soldiers that is organized to provide its own support. These support units are classified as follows: Combat Support (CS) is Signal (Communications and Computers), Military Police, Intelligence, Etc. Combat Service Support (CSS) is Logistics (Supply Maintenance, and Transportation), Ordinance, etc. This extra "slice" of support equipment is that portion of CS/CSS that supports the brigade, but that would normally be assigned to a division.

From left to right, the brigade in the Figure 16 consists of the following units according to TSP158-A-3000 [TSP98]:

- Two armor battalions
- An engineer battalion
- A mechanized infantry battalion
- A scout company
- A signal section
- A field artillery battalion
- A military police section
- An air defense battery
- A headquarters company (HHC)
- A support battalion (medical, maintenance, etc.)
- A military intelligence company

A pure armor brigade would have three armored battalions. This brigade has two armor battalions and one mechanized infantry battalion; therefore, it is *task organized* according to HQDA [HQD92].

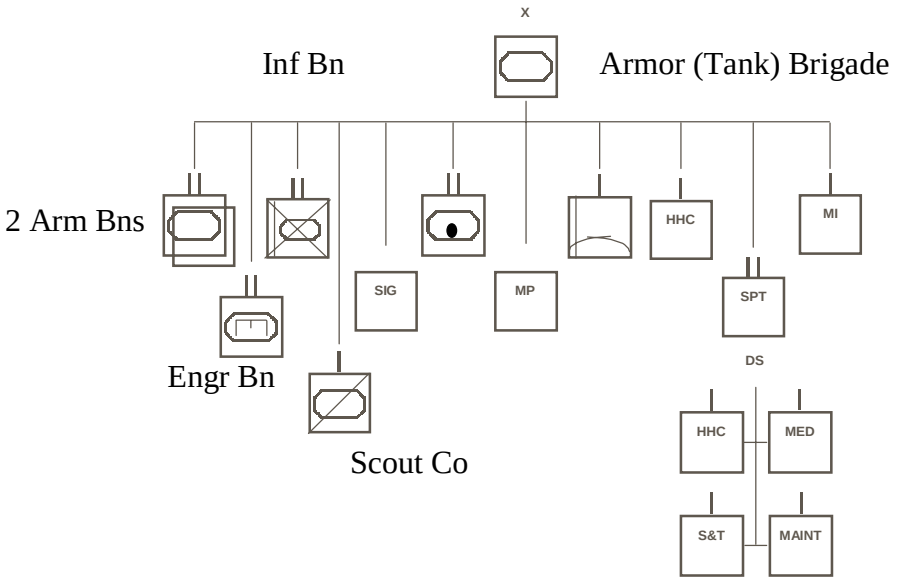


Figure 16 Separate Brigade Organization - HQDA [HQD96a]

Again, for the purpose of this research, assume that each ship in the APS afloat set has a Battalion Task Force (BNTF) and several support unit slices. For example, if the entire collection of ships is represented by Figure 16, one ship may have an armor battalion, an engineer company from the engineer battalion, an artillery battery from the artillery battalion, and platoons or sections from each of the other support units. Each BNTF ship is either an armor or an infantry battalion that has been task organized. In addition to the maneuver companies, other elements, such as a field artillery battery, are added so that the BNTF may have additional mission flexibility. Additional units on the ship represent the "slice" of Combat Support and Combat Service Support (CS/CSS) that the battalion would normally receive from higher headquarters. These are the Brigade (BDE) CS/CSS slice, the Echelons Above Division (EAD) CS/CSS slice, and the Corps / Theater Base (CTB) slices of support equipment.

The traditional combat mission *task organization* is further complicated by the requirement to have the units on ships. In addition to being organized for the mission, the units have two additional constraints. (1) If one or more of the ships is lost, Army

planners want the surviving ship to be "complete" in that it can be organized and sustained for a contingency. (2) The logisticians who load ships want the ships to be completely "full." They wish to completely fill the ship without absolute regard to the unit composition. Given these competing constraints, the *task organization* of the ship battalion is typically adjusted during the ship load-planning process to ensure that the unit on the ship is self-contained and to ensure that the ship is packed to capacity.

All units are rated each month based upon their readiness for war. For many units, these ratings are classified at the level of SECRET or TOP SECRET. The calculation of the equipment readiness ratings is complicated because of the task organization. The required readiness calculations, which must be reported to the Army leadership, are normally based upon the authorized units and not upon the task organization. This is further complicated in the afloat set because parts of units are permanently task organized and physically located on several ships. This readiness ratings data must be calculated across several dimensions at each echelon.

An analyst may choose to examine the summarized readiness data from two common dimensions. The first is by authorized unit and the second is by location. Location corresponds to a single ship and a single battalion task force. This common requirement complicates the data required in the example domain for this research.

The data that describes the equipment location, authorization, on-hand quantities, and other attributes must be gathered from several legacy data sources in order to determine the readiness of the equipment on one or more of the ships. Over a dozen unclassified samples of representative Army War Reserve data sources were obtained for this research. The data and schema descriptions were obtained from both the Logistics Integration Agency and the Army Strategic and Advanced Computing Center Data Warehouse.

Kumara [KUM99] and Huhns [HUH99] have proposed agent-based approaches to access distributed data in similar logistics problems. This research uses mediator agents to access the heterogeneous legacy data sources and visual indirect manipulation interfaces

to enhance the coherence of the information to provide answers to representative tasks. The next section briefly examines agent-based approaches to set the stage for the methods used to access the data in this research.

2.5.2 Software Agents

> If God had an agent, the world wouldn't be built yet. It'd only be about Thursday. - Reynolds [REY91]

The term "agent" has appeared in several computing literature fields from artificial intelligence to user interfaces. Definitions of agency range from those that include nearly every computer program to those that require more autonomy that many human beings exhibit. Franklin and Graesser [FRA97b], provide an overview in their attempt to answer, "Is it an Agent, or just a Program?" Wooldridge and Jennings [WOO95b], offer a definition of weak agency as a computer system that enjoys the properties of autonomy, social ability, reactivity, and proactivity. At the other end of the spectrum, Shoham [SHO93] advocates a strong agency that consists of mental components such as beliefs, capabilities, choices, and commitments.

A generic agent system architecture is shown in Figure 17. In that figure, the *User Interface Agents, Mediator Agents, and Data* on the left side of the figure can represent the research system. Additional brokers, ontology agents and data sources, and applications programs are available in the agent development environment. Many of these agents are not absolutely necessary for remote usability research of agent-mediated legacy data.

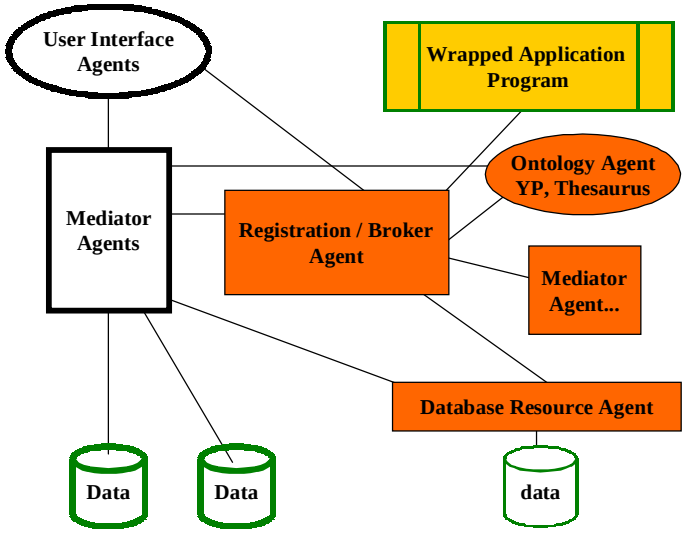


Figure 17 Generic Agent Architecture - Huhns and Singh [HUH98a]

At the most basic level, an agent is a process that communicates with and performs information preparation and exchange on behalf of a client or server. This communication may be with another agent or user or non-agent and the process may be automatic. Russell and Norvig [RUS95] argue that strict definitions that sharply divide agents from non-agents are less useful that the notion of an agent as a tool for analyzing systems. This weak agency notion is the one preferred by this research.

Despite the various roots of agent research, much research motivation concerns interoperability among users, information sources, and tools. Today's interest in agents results from the convergence of earlier work in many fields. This convergence may be seen, at least in part, as trend towards enabling increased interoperability. Interoperability between legacy and allied systems is especially important in the defense domain. Several major areas of defense-sponsored research are currently ongoing in the interoperability field. These research areas, SHADE, LISI, and KSE are summarized below.

A portion of the Defense Information Infrastructure, DII, is concerned with the ability to share data among heterogeneous systems. The resulting research, called SHADE (SHAred Data Environment) is designed to provide components such as database runtime tools to allow access to data in specified database systems. SHADE also provides developer tools to support the development of sharable database segments according to Hartley [HAR96a].

In Joint Vision 2010, the United States Joint Chiefs of Staff, JCS [JCS98c], state:

> *Forces harnessing the capabilities potentially available from this (C4ISR) system-of-systems will gain dominant battlespace awareness, an interactive 'picture' which will yield much more accurate assessments of friendly and enemy operations within the area of interest. Although this will not eliminate the fog of war, dominant battlespace awareness will improve situational awareness, decrease response time, and make the battlespace considerably more transparent to those who achieve it.*

In order to achieve this dominant battlespace awareness, the fragmented C4ISR (Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance) communities must become interoperable according to C4ISR-AWG [C4I98a]. A model for assessing interoperability was developed and described in C4ISR-AWG [C4I98b]. The LISI (Levels of Information Systems Interoperability) model describes the interoperability attributes of information systems. The levels range from the isolated level zero with manual entry and re-entry of data to the enterprise level four with universal multinational, interactive access to many models. This research enables interoperability that is associated with levels 2 and 3 of this model as shown in Figure 18.

Finally, the Defense Advanced Research Project Agency, DARPA[1] has funded several interoperability efforts. These efforts include: I3 (Intelligent Integration of Information) as described in DARPA [DAR98], and the KSE (Knowledge Sharing Effort) described in KSE [KSE98]. These efforts seek to increase interoperability through the integration and sharing of data and knowledge sources.

[1] Sometimes known as the Advanced Research Project Agency, ARPA.

Level	Description
4 – Enterprise	Global information and application sharing. Interactive collaboration
3 – Domain	Shared Data. Separate applications. Common operational picture.
2 – Functional	Heterogeneous product exchange. Separate data and applications.
1 – Connected	Homogeneous product exchange. Electronic connection.
0 – Isolated	Manual product gateway. Disk, and paper copy exchange

Figure 18 Levels of IS Interoperability - C4ISR-AWG [C4I98b]

An example of the integration and sharing of data and knowledge sources in distributed web applications is offered to highlight several of the capabilities and limitations of interoperability agents as described in Huhns [HUH98c].

Suppose an Army unit wants to order supplies over the web, debit an account, and track delivery. The application must record a requisition in a requisition database, debit the account, send the order to the shipping agency, receive a tracking code, and update an inventory database. Two possible problems that may be caused by an interrupted transaction include shipping the order without debiting the account and debiting the account without entering or shipping the order.

In a closed environment, several commercial solutions exist. Transaction processing (TP) monitors, such as IBM's CICS, Transarc's Encina, and BEA System's Tuxedo ensure that either all or none of the steps are completed, and that systems eventually reach a consistent state. However, if the user is disconnected just after she selects "OK,"

or the line is disconnected, the order may or may not succeed. In this case, the TP monitor may not be able to get the user into a consistent state as described in Cross [CRO97].

In an open environment, some system of more intelligent components, called agents, must cooperate to solve the problems experienced by the Army unit attempting to order supplies. For example, a server application could send email about account problems, or detect duplicate transactions as shown in Duncan [DUN98]. A downloaded Java applet could synchronize with a server after a broken connection was reestablished as discussed in Frost [FRO97]; Watson [WAT97]. The applet could recover the transaction and communicate with the server or directly with server objects via other agents. If there are too many orders to process synchronously, they could be put in a message queue, managed by another agent server that guarantees message delivery or failure notification. The ordering unit could be notified by email when the transaction is complete. Each of these solutions requires some system of agents as described by Decker, Sycara, and Williamson [DEC97a].

The Interactive Maryland Platform for Agents Collaborating Together (IMPACT) platform provides such an agent environment. IMPACT defines a set of registration, yellow pages (yp), thesaurus, type, and interface infrastructure servers in a hybrid multiagent architecture described by Subrahmanian et al. [SUB00]. The IMPACT architecture provides an agent service description language and algorithms for wrapping program code to allow agent communication. A complete specification of this software infrastructure may be found in Arisha et al. [ARI98].

To deploy a new agent on the network, the agent's creator registers its services with the registration server. The registration service Java interface communicates with the ontology servers (thesaurus, yp, and type) allowing the creator to hierarchically browse similar description words, service descriptions, and types, respectively. When a registration is complete, noun, verb, type, and agent-table hierarchy data structures are updated.

Agents may ask the yp server to identify other agents that furnish particular services. The yp server performs a nearest neighbor search to identify likely candidates and then the requesting agent may contact the candidates directly for the particulars. In both the search and the agent-to-agent communications, the infrastructure servers allow the agents to share the same ontology and to communicate about a domain without necessarily operating on a globally shared view. In this way, the IMPACT platform adds robustness to the computing environment. Agents may discover other agents with no fixed network address, and if one agent is missing, they may find another one that suffices.

Several researchers are conducting extensive work with the IMPACT platform as described in Subrahmanian [SUB99a]; [SUB99b]. This dissertation research concentrated on customizing and developing two stable and essential components of many agent architectures: the mediation server (HERMES) and the interface agents (JIMI and VITAMIN).

In the IMPACT architecture, HERMES serves as the wrapper and the link to arbitrary data sources and external programs. HERMES is the current data source for IMPACT. One of the main problems in accessing distributed agents is the connection problem as shown in Davis and Smith [DAV83]. The connection problem refers to the issue of locating the agent with the information or capabilities that are required according to the requesters' preferences as discussed in Decker et al. [DEC97b]. The location problem is not the focus of this research; therefore, the interface agents are not required to interact with the IMPACT servers to identify the location of the HERMES mediator agent used for this research.

Mediators, pioneered by Wiederhold [WIE92], are the central agents that perform the integration and interoperability functions as shown in Figure 19. Huhns and Singh [HUH98a] describe mediators as agents for information-rich environments. The central mediator agent offers solutions to the interoperability challenges discussed in the previous section. There are many challenges with respect to visualizing networked heterogeneous data, relationships, and structure according to Adali et al. [ADA96a]. These challenges are similar to those related to other very large data sources whether in

data ware houses or mediated from raw legacy files as shown in Keim, Kriegel, and Seidl [KEI94] and Goldstein and Roth [GOL94].

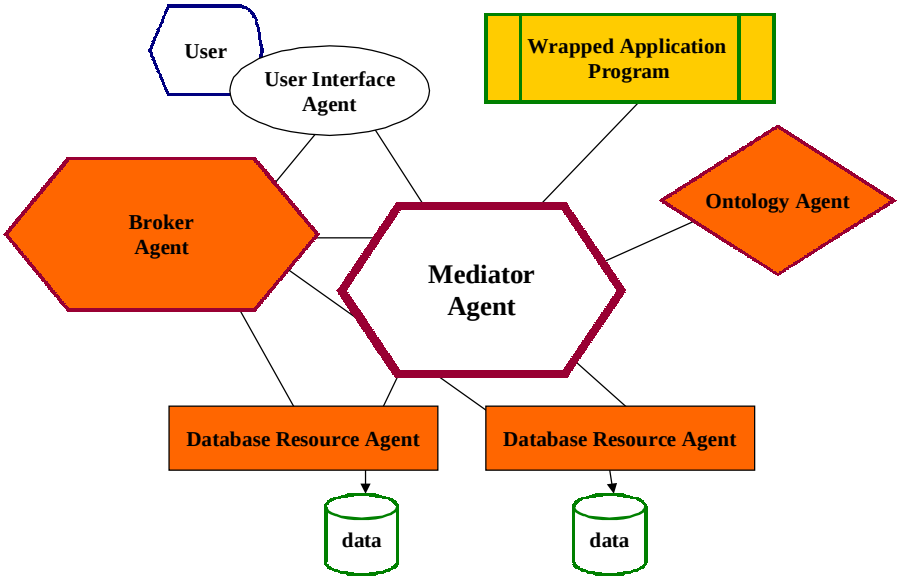


Figure 19 Central Mediator Agent - Huhns [HUH98c]

In order to integrate and visualize these disparate data sources, there exist two obvious alternatives: (1) Integrating the data into a single source such as a data warehouse; (2) Allow the data sources to remain with their originator, and create a system to retrieve the data as needed. Combining the data into a single source has the advantage of centralizing the data and providing users with a common interface. For analysis of static snapshots of data, On-Line Analytical Processing (OLAP) is often appropriate. OLAP usually involves large amounts of diverse data aggregated into a data warehouse as described in Bontempo and Zagelow [BON98c] and Sutter [SUT98]. The OLAP software provides an interface for users to transform raw data and interactively examine the results in various dimensions to look for patterns as discussed in Navin [NAV98], and Chaudhuri and Dayal [CHA97b].

However, disadvantages of importing data into central warehouses include failures due to single-source centralization. Another issue is keeping the imported data replicated, current, and consistent with the original source. In addition, many data warehouses are built on a relational model of data. Relational structures and operators are not efficient for certain data types such as multimedia, geographic, and non-relational legacy data. A single data structure may not efficiently support all required functions, for example, storing GIS map data in a relational database is very inefficient according to Adali and Subrahmanian [ADA96c].

In addition, legal issues of proprietary data and classification of aggregated data sources may prevent users from gathering all required data in one location. For instance, individual query results may not be classified or proprietary. In fact, a data source provider may be in the business of providing answers to queries. However, the entire collection of data may be classified or proprietary.

Unit readiness data, the functional domain of this research, is classified when aggregated according to HQDA [HQD97a]. Additionally, for this research, different organizations are responsible for and separately maintain the various sets of data legacy. This research explores the second alternative of allowing the data sources to remain with the originator and creating a mediator agent system to retrieve the data as needed.

Mediator agents characteristically express methods to resolve conflicts, unify mismatches in measurement units, and generate sophisticated conclusions based on information contained in a wide variety of data structures. Mediators are typically built in two steps, package integration, and semantic integration. Package integration provides access to functionalities supported by database or package. Semantic integration merges the results of applying these functionalities as shown in Buneman, Raschid, and Ullman [BUN97] and Rosenthal et al. [ROS97].

Mediator features usually include minimal client requirements since data remains in the original site, efficiency since the data remains stored in native data structures, and quick integration. Interoperability is hastened because data is accessed through the predefined,

legacy software that is used to manipulate the data as shown in Arens et al. [ARE94] and Ashish [ASH98]. This research uses the HERMES mediator agent.

HERMES defines a platform for building mediators that semantically integrate different and possibly heterogeneous information sources and reasoning systems. HERMES mediator-agents employ annotated logic-based rule sets that define precise domain function execution over target data sources. HERMES domain modules encode the actual conduit through which the system accesses native data files according to Subrahmanian et al. [SUB97a].

The HERMES mediator agent is able to integrate several types of data sources and planning tools including flat files, Ingress, Oracle, Objectstore, image data, GIS data, video data, Army route planner, a face recognition system, and operations research software. In addition to Oracle data sources, custom HERMES domains for several legacy data sources were created for this research.

2.6 Open Issues

From the preceding sections, emerge several open issues in computer science. First is the need for remote empirical evaluations of the usability of agent-mediated heterogeneous legacy data sources in a specific, well-defined context. The environment addressed by this research is that of the increasingly common and complex distributed heterogeneous legacy data sources. The precise environmental context is that of Army War Reserve equipment readiness. The research problem is to determine whether an enhanced indirect manipulation interface system can add coherence to agent-mediated legacy data for users performing representative tasks.

Researchers have shown that a method to enable this coherence enhancement is to improve the user interface system. In order to accomplish this, new interface systems must designed, built, and tested. To test the efficacy of a new enhanced system, it must be compared with an unimproved traditional legacy interface system. Traditional and enhanced IMISs were constructed for this research. Several enhanced visual cues were identified in the literature and reviewed in this chapter. These include improved data

selection instead of data entry, improved visual clarity with familiar color effects, and improved representation with an aggregating tree structure.

A common vocabulary and methodology is required to describe the usability metrics and the interaction metaphor in order to compare these two interface systems. The concept of coherence and the interaction metaphor of indirect manipulation are defined here to meet this requirement. Other researchers have defined concise usability metrics, but these have been designed for "quick and dirty" tests instead of statistical analyses. Coherence is used as both a theoretical framework for a characteristic of the interface system and as a concise implementation metric for statistical analysis of usability.

Finally, the opportunity to design a straightforward experimental environment to enable remote testing was identified. Such a system may allow collection of experimental data from which statistical analysis may be conducted. Such an innovative remote evaluation system may also provide more realistic results than laboratory evaluation. It must enable the four important usability factors, effectiveness, efficiency, confidence, and satisfaction to be recorded during a remote usability evaluation. A remote evaluation system and methodology was constructed to allow the direct comparison of the indirect manipulation interface systems. This empirical evaluation environment also addresses the need for more experimentation in computer science.

This research found no empirical evaluations of usability in the context of an agent environment to legacy data sources. In networked information systems where heterogeneous legacy data is distributed and network access may be slow, this research uses a coherent, indirect manipulation interface system (IMIS) to access a mediator agent system.

Chapter Three, *Methodology*, presents a detailed description of the research systems developed and the experimental design constructed to test these theories.

CHAPTER 3: METHODOLOGY

> Traditional scientific method has always been at the very best, 20-20 hindsight. It's good for seeing where you've been. - Pirsig [PIR74]

Chapter Two revealed several open issues in usability and human computer interaction research. To address these issues, this chapter describes three prototype systems that were constructed to gather empirical support for the theoretical concepts put forth in chapter two. These systems include the JIMI and VITAMIN indirect manipulation interface systems (IMIS) and the JUICE remote evaluation system.

Many organizations require quick and efficient visualization and manipulation of legacy data sources with diverse data structures. The methodologies developed for this dissertation are valuable in many circumstances. To test the systems with actual data, a military logistics example was chosen as representative of many challenging organizational domains. To illustrate, consider the problem of developing a plan to deliver equipment for an Army battalion task force in support of a contingency operation. The equipment must be delivered from stocks that are pre-positioned aboard various ships around the world. The creation of such a plan requires consolidation of information from many distributed, heterogeneous data sources.

These data sources include huge legacy files of equipment authorizations, logistics supplies, parts transactions, unit authorizations, task organizations, ship-load plans, and various location details. A specific challenge addressed by this research is to provide efficient access to this data and to display the readiness of the pre-positioned equipment in a coherent manner is. The goals of this research were to determine whether novel interface systems can provide improved coherence to this agent-mediated legacy data and to develop a unique remote usability evaluation system in which to conduct the empirical comparisons.

This research developed a methodology to improve user querying of agent-mediated legacy data. In addition, a remote usability interface evaluation system was developed to

empirically study the efficacy of the methodology. The research experiment examines whether an indirect manipulation system can enhance user querying of agent-mediated heterogeneous legacy data sources. Specifically, the research address whether such an enhanced coherence interface can assist minimally trained users in answering common representative task questions.

Two interface systems were designed to examine the research methodology. The Java Indirect Manipulation Interface system, JIMI, was designed to emulate a traditional query interface. The Visual Interface to Agent Mediated Information Networks system, VITAMIN, was designed with enhancements to add coherence to the user querying of the agent-mediated legacy data. Interface systems enhancements identified in chapter two are not designed into JIMI. These enhanced visual cues include improved data selection instead of data entry, improved visual clarity with familiar color effects, and improved representation with an aggregating tree structure. Both of these indirect manipulation interface systems construct query actions that are sent to a mediator agent that, in turn, accesses the heterogeneous legacy data sources.

A third system, the Java Usability Interface Comparison and Evaluation (JUICE) system was designed to enable the remote usability evaluation experiment, which compares the coherence of VITAMIN with JIMI.

The remainder of this chapter details the research systems and hypotheses, experimental design, and ethical considerations.

3.1 Research Systems and Hypotheses

> The nation that will insist upon drawing a broad line of demarcation between the fighting man and the thinking man is liable to find its fighting done by fools and its thinking by cowards. - Butler [BUT74]

This research used a quantitative paradigm methodology for the empirical usability testing of the VITAMIN system and the JIMI system. Within the quantitative paradigm, a randomized or true experiment is used to show a causal relationship as opposed to a quasi experiment or non-experiment that may merely indicate a correlation. The remote,

task based, usability experiment determines statistically whether the VITAMIN system is superior to the JIMI system.

A formative evaluation of the JUICE system was conducted during the pilot studies. The evaluation design was a one-shot survey. This non-experiment gathered formative and descriptive data that was used to improve the design of JUICE. Open-ended design questions were also provided for JIMI and VITAMIN to add qualitative richness to the quantitative data. The next three sections describe VITAMIN, JIMI, and JUICE in detail.

3.1.1 Visual Interface to Agent Mediated Information Networks (VITAMIN) System

The VITAMIN system described provides comprehensive presentation and navigation of Heterogeneous AWR Legacy Data Sources (HALDS). Specifically, the system enables analysts to display and navigate the go-to-war readiness of military equipment prepositioned aboard ships directly from the original data sources. This research uses unclassified subsets of the actual classified data.

On the VITAMIN system screen, the user selects more detail or less detail about the materialized views. Pointing and selecting a color-coded "plus" or a "minus" button formulates a query. The overview frames present a directory style tree list of aggregated data. These frames adjust as the user expects such a list to adjust as shown in Figure 20.

If the user chooses location, then a query can be generated to retrieve selected data sets to determine the available locations. The aggregated list of locations available is returned, and the readiness and the underlying authorized and on-hand equipment quantities are presented in a simple visual query panel. This high level aggregated summary may answer the user's task question or this information may indicate the need to further disaggregate one or more locations to investigate an anomaly. If the user chooses, she can then click on one of the locations to view information at the next lowest level of the location. Once again, this information is taken directly from the mediated data and may not be known to the user or the developer *a-priori*.

Conceptually, the client indirect manipulation interface system (IMIS), VITAMIN, implements a forest of multi-branch HERMES query trees which enable logistic planners

to drill-down into or disaggregate all necessary data paths while shielding users from the underlying HERMES action syntax. The multi-tree, multi-branch aspect allows the user to visualize the data from many different perspectives. For example, several attributes may characterize a particular data set uniquely. The client interface allows the operator to select a set of data visually by one set of attributes and then examine the data by another view as described by Martin, Cheyer, and Moran [MAR99b].

Architecturally, VITAMIN exists as an HTML-embedded Java applet that communicates with the remote HERMES server via standard TCP/IP socket protocols. This allows logistic planners at multiple locations, working often from diverse operating systems and platforms, to visualize the required Army readiness data using standard Java-enabled web browsers.

The VITAMIN System used for this research is in the *VITAMINL* (vitamin for logistics) Java package. The *VITAMINL* Java class encodes the query tree. Each query can map to zero, or more sub-queries. As users select to drill-down, the client interface maps data from the parent query answer to its child sub-query masks. For example, suppose a user wishes to see what equipment is located on the notional supply ship *Alexandria.* To drill into the *Alexandria* data the client interface maps data (the location selection) from the *loc_totals* query answer: loc_totals (3, *Alexandria*, 9510, 8471, 89.0747):1 onto its child query nodes. In this case, it generates two sub-queries:

```
loc_ERC_totals (D_Status, Alexandria, SzERC, D_AuthQty, D_OnHand, D_Percent):1.0

loc_force_totals (D_Status, Alexandria, SzFrc, D_AuthQty, D_Onhand, D_Percent):1.0.
```

These two queries, one in the “ERC” dimension and one in the “force” dimension, ultimately render the six answers appearing indented beneath the *Alexandria* location data. ERC (Equipment Requirement Code) encodes the relative importance of the equipment type, and the force maps to the subordinate units.

The VITAMIN system presents a visual display of the mediated data hierarchy as shown in Figure 20.

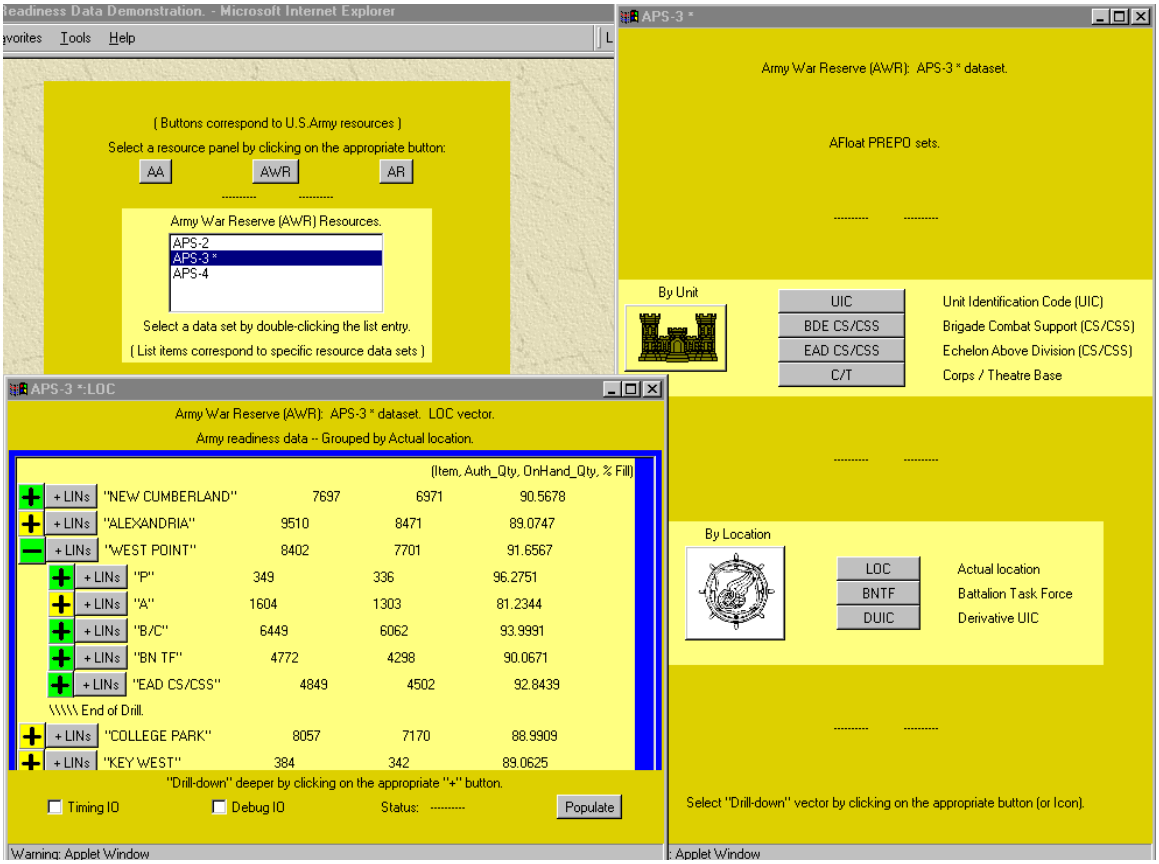


Figure 20 VITAMIN Screen Capture

In the top left panel of Figure 20, the user has selected the mediator for the Army War Reserve data sources and then within those data sources, the user has selected APS-3, the afloat set. VITAMIN then offers the panel on the right for the user to choose dimension from which to investigate the data. The user chooses the By Location / Actual Location dimension button and is presented with the panel on the lower left. In this screen shot, the user has also drilled into the Item / Location "West Point" to see the readiness by both types of equipment (ERC P, A, B/C) and subordinate units (BN TF, EAC CS/CSS).

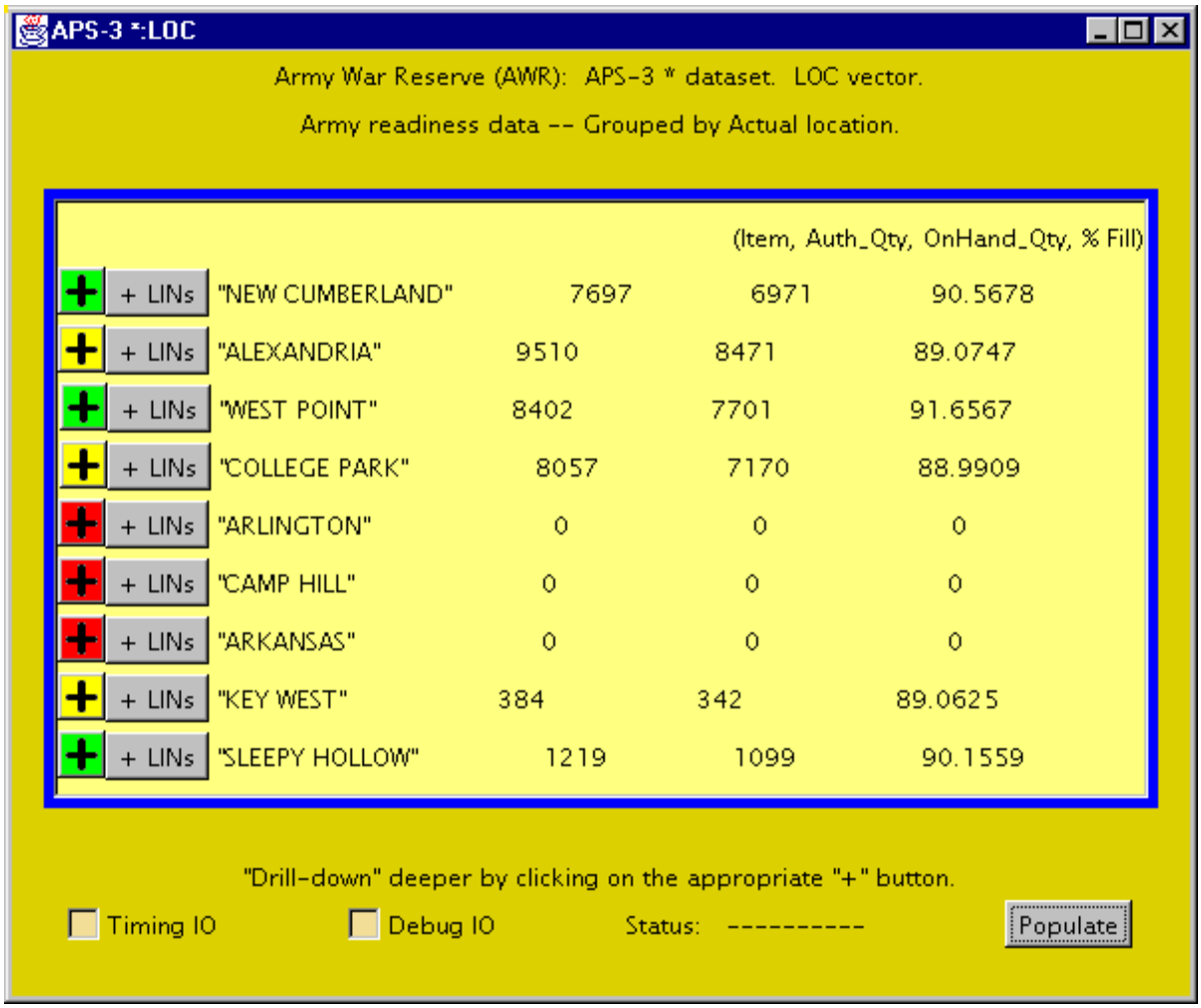


Figure 21 Equipment Readiness On Fictitious Supply Ships

In the VITAMIN system, the user uses a mouse to select and specify each query. Then the system responds with the requested data, which is presented as one more layer of data disaggregation. Figure 21 through Figure 23 show the screens generated as a user drilled first into the *Alexandria* readiness data, and then specified *Alexandria* Battalion Task force (BNTF) readiness data.

Recall that the *actual location* is the same as the *ship name* for Army Prepositioned Stocks-3 (APS-3) afloat stocks. The figure also illustrates two dimensions of the *Alexandria* readiness (1) by Equipment Requirement Code (ERC): P, A, B/C and (2) by next lower unit: BNTF, Brigade (BDE) and Echelons Above Division (EAD) Combat Support and Combat Service Support (CS/CSS) as discussed above.

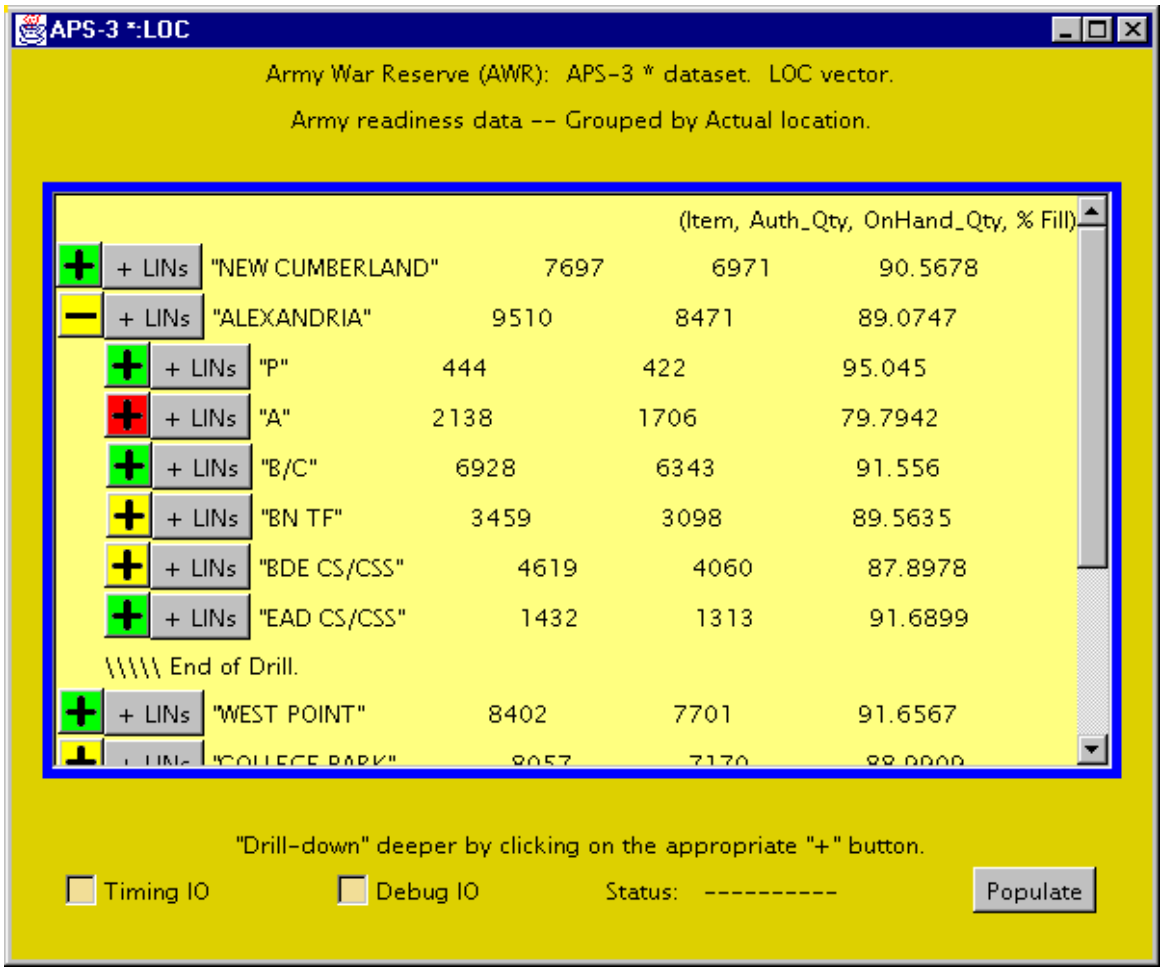


Figure 22 A Drill-Down Of The Fictitious Supply Ship Alexandria

Note the plus (+) and minus (-) drill control button color-coding currently represents a straightforward, unclassified, percentage based readiness level indicator implemented for this research. The colors correspond to typical color usage in the readiness domain according to White [WHI98]. The colors have the following connotations: Green (medium) >= 90% > Yellow (light) => 80% > Red (dark) < 80%. Both the actual analysts and the test subjects are familiar with this color usage where green is the best and red is the worst. A production implementation would render these colorings based on a much more complex and classified user specified data function call as described in HQDA [HQD97a], and Smith and Schafer [SMI98a].

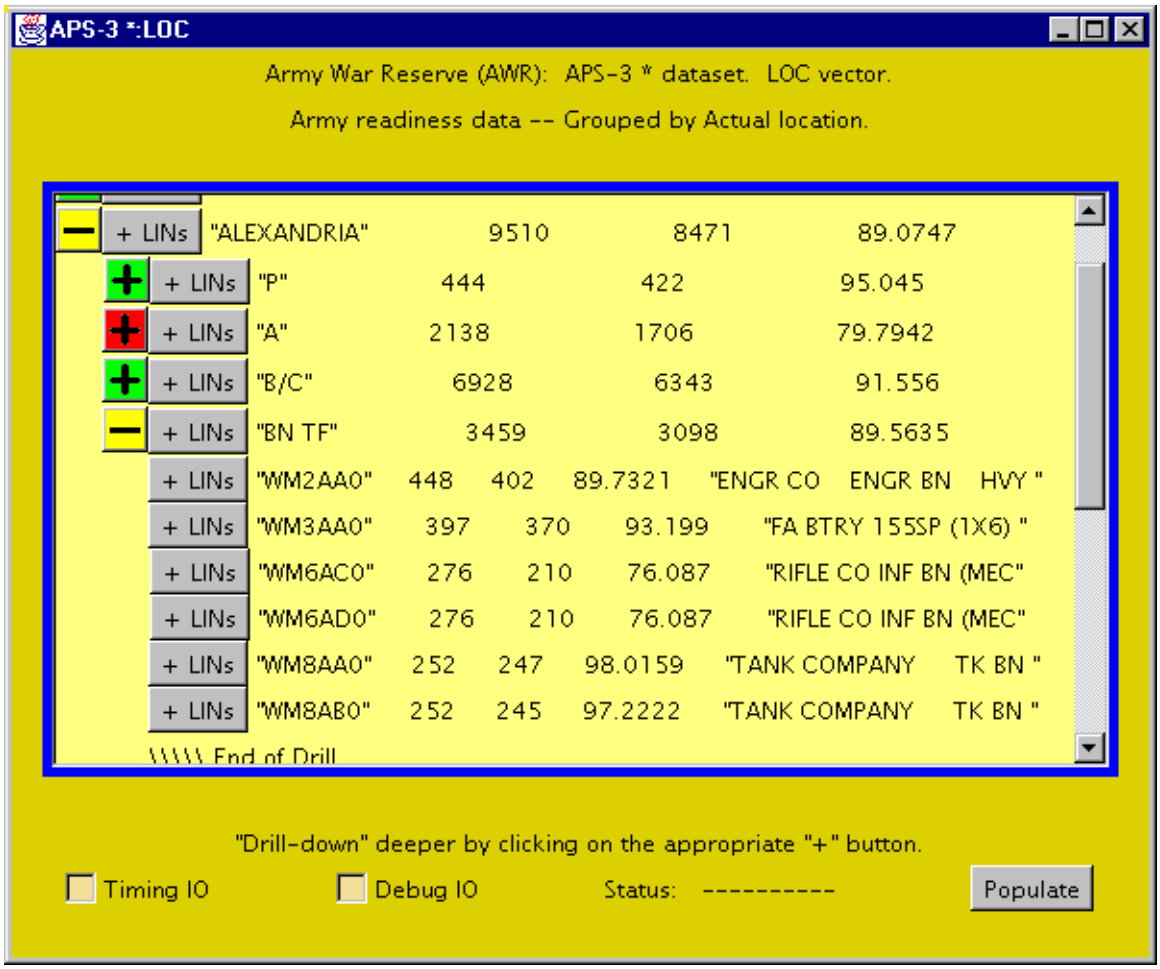


Figure 23 A Drill-Down To The Unit Level

Pointing and selecting or "clicking" on a table entry plus ("+") button causes the system to disaggregate one level deeper into the data. The next level's data then appears indented in the table beneath its parent entry and the plus button changes mode to minus ("-"). This action is reversible. Subsequently clicking on a minus button for any level causes that aggregation path to collapse.

VITAMIN includes the following theoretical enhancements identified in chapter two as compared to the traditional JIMI system described in the next section: (1) pointing and selecting instead of using command line data entry; (2) summarizing tree for data aggregation; and (3) color-coded query buttons.

3.1.2 The Java Indirect Manipulation Interface (JIMI) System

The Java Indirect Manipulation Interface (JIMI) presents a much simpler and more traditional interface to the user. Users are presented with queries which they can instantiate and customize before execution. JIMI also performs queries on a mediated Heterogeneous AWR Legacy Data Sources (HALDS). The output of a query is the list of units and readiness matching the specifications of the query. Vertical and horizontal scroll-bars can be used for scanning the list.

Figure 24 shows a screen capture of JIMI. After logging in, the user selects the mediator file (left panel), then the predicate (upper right) and when the query is executed, the result is displayed in the panel on the lower right.

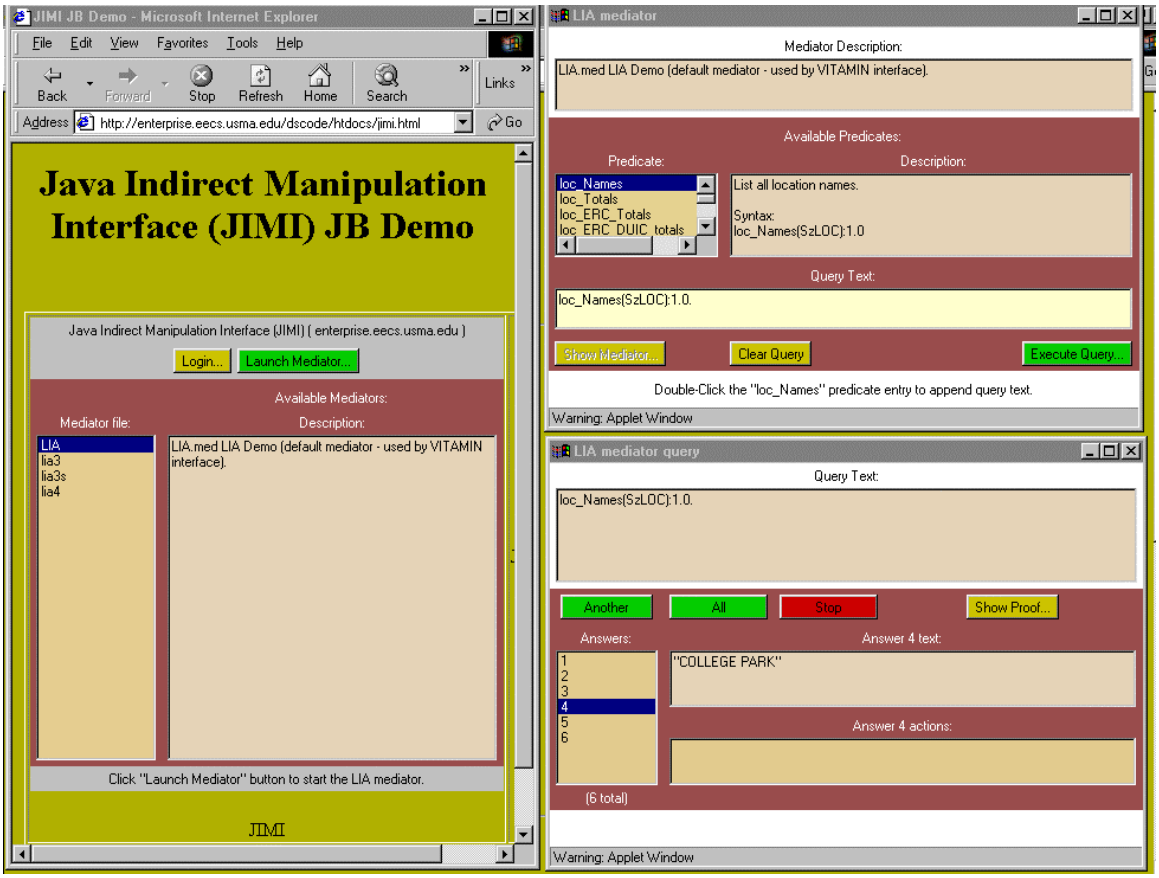


Figure 24 JIMI Screen Capture

JIMI was designed to represent the current legacy querying system interface used by actual users as described by White [WHI98]. In order to evaluate the interface system differences properly, JIMI differs from the traditional legacy interface system in two ways. First, JIMI uses HERMES to access the data instead of using a single snapshot homogeneous database. Secondly, JIMI displays example query predicate actions to enable non-expert subjects to participate in the experiments.

JIMI and VITAMIN both use the HERMES agent to access the legacy data in order to provide an accurate comparison of the interface systems. Users should generate identical query actions from both the VITAMIN and JIMI indirect manipulation interface systems to answer the same task. The query actions that JIMI sends to the HERMES mediator agent are identical to those sent by VITAMIN. Therefore, there is no interference from HERMES to confound or bias the dissertation experiments.

Like VITAMIN, JIMI is a Java applet as well. For the purposes of this research, both of these Java applets are embedded inside another Java applet. This third system, JUICE, is the remote usability evaluation system that presents the representative tasks and records the experimental results. JUICE is described in more detail in the next section.

3.1.3 Java Usability Interface Comparison and Evaluation (JUICE) System

A screen capture of the JUICE remote testing system is provided in Figure 25.

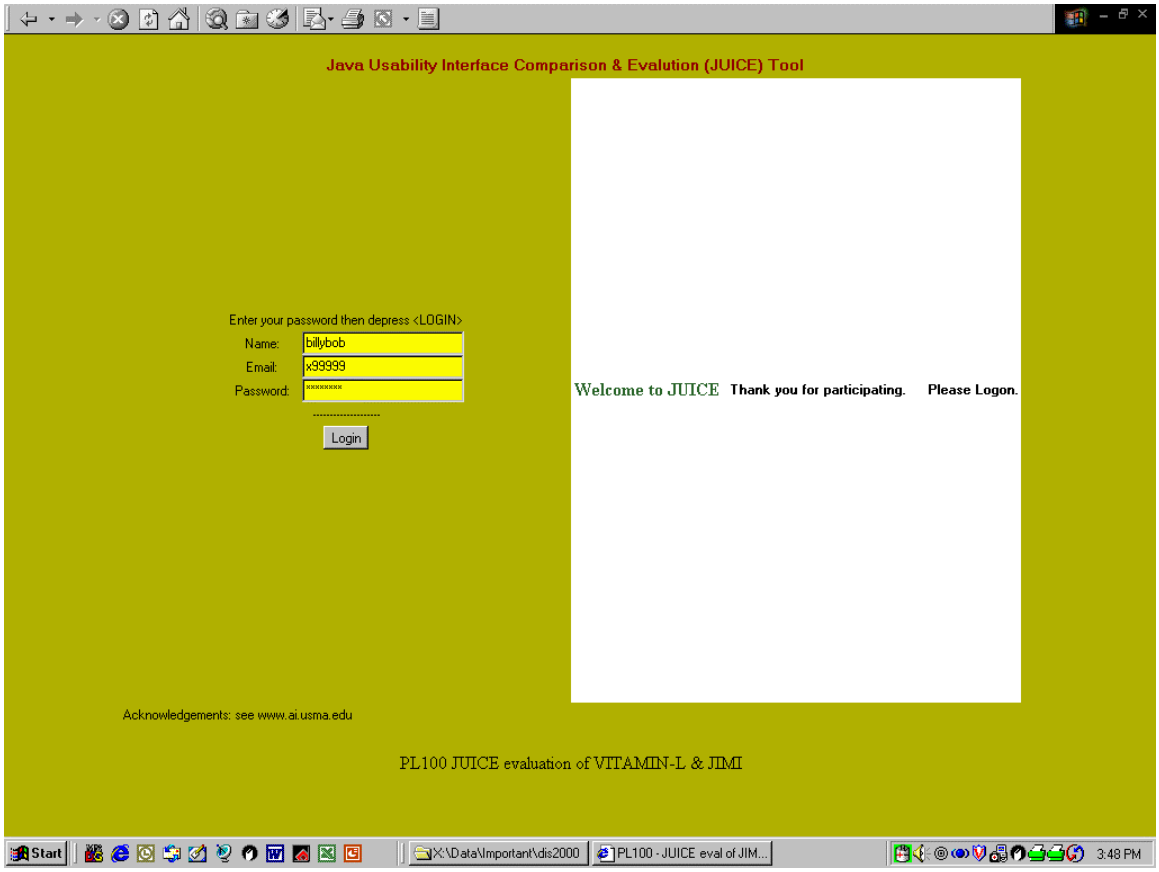


Figure 25 JUICE logon

The Java Usability Interface Comparison and Evaluation (JUICE) system prototype is a remote usability-testing system designed to allow comparison and evaluation of two different Java indirect manipulation interface systems (IMIS). The JUICE system was designed to evaluate the IMIS used in this research, VITAMIN, and JUICE.

JUICE meets the need for an innovative remote usability system articulated in chapter two. It enables the four important usability factors, effectiveness, efficiency, confidence, and satisfaction to be recorded during a remote usability evaluation. JUICE does not require embedded metering code; rather, the Java applications to be compared are embedded into JUICE, the evaluation system.

JUICE enables a randomly assigned IMIS treatment for a within-subject counterbalanced experiment. After logging into JUICE, subjects are presented with surveys, treatments,

and tasks. JUICE collects all of the subject's actions and responses on a central server for later analysis.

Details of a complete JUICE session and functionality are described in the *Research Procedures* portion of the *Experimental Design* section later in this chapter. The next section describes the mediator agent actions that both indirect manipulation interface systems, JIMI and VITAMIN, send to the HERMES mediator agent in order to access the distributed heterogeneous legacy data sources.

3.1.4 Mediator Agent Actions

The HERMES mediator agent used for this research mediates among five representative legacy data sources shown in Table 1.

Data Source	**Partial Contents**
aps_loc	Army Prepositioned Stock locations
sb700_20_H	Equipment descriptions
equipRU	Actual equipment quantities on ships
aps_frc	Force Structure, Unit hierarchies
wm_moc	Authorized unit equipment quantities

Table 1 Data Sources

Once the data was amassed, the next step was to create the necessary data interfaces. This required implementing the HERMES mediator agent domains for the legacy files. For example, the wm_moc file required the creation of a LOGTAADS (Logistics - The Army Authorization Document System) HERMES mediator agent domain. The LOGTAADS data set exists in "flat-file" form but encompasses four distinct record types, typical of many legacy COBOL data files. It is a single-file, multi-table, relational database that lacks an application-programming interface (API). Access to this legacy data required writing the LOGTAADS domain (a HERMES software interface) that defined basic relational database operations over the native legacy LOGTAADS tables.

HERMES mediator files (mediators) list declarations and rules for accessing and combining diverse software and data. The mediators map query action requests onto data sources through domain function calls such as to the LOGTAADS domain. Mediators generally correspond to applications. For example, HERMES mediators have been constructed for applications such as inventory management, travel planning, and networked banking systems.

The mediators for the AWR planning system define the query action predicates, data sources, rule bodies, and security conditions necessary to instantiate the appearance of a unified homogeneous database. Again, it is important to emphasize that the mediator must integrate disparate U.S. Army data sources by invoking required sequences of LOGTAADS domain function calls with function calls to other domains.

The development environment used to create and implement these domain function calls as well as the development environments required for the accompanying research systems are described in the next section.

3.1.5 Development Environment

The software development environment and tools are as follows:

- Java: JIMI, VITAMIN, and JUICE are written in Java. They run in browsers.
- Front Page: The experimental web site was developed in Front Page.
- Oracle: The relational tables are stored in an oracle database on an NT server.
- Apache: The web server is Apache running on a Silicon Graphics Inc. (SGI) server.
- HERMES: The mediator agent runs on a Solaris server.

The VITAMIN system interface was developed in Java. Development was accomplished first with the SUN Java Development Kit (JDK), then with Microsoft Visual J++ (VJ++) and finally with Inprise (Borland) JBuilder. The JIMI system was developed in Java

using Inprise JBuilder. JUICE was developed with JBuilder and MS VJ++. HERMES and the mediator functional domains are written in c using the Solaris development environment.

Now that the implementation and development of the research systems and actions have been detailed, the research problem, research hypotheses, and null hypotheses are addressed in the next three sections.

3.1.6 Research Problem

Can an enhanced indirect manipulation interface system add coherence to agent-mediated legacy data for users performing representative tasks?

This research proposes to show statistically significant improvements in coherence as measured by the number of *correctly* answered tasks, user *confidence* in task answers, time to answer tasks (*speed*), and, user *satisfaction* with the interface system (C^2S^2). For succinctness, the following details are specified for each hypothesis.

- Results are statistically significant
- Subjects are members of the USMA class of 2003 (Pilot study were '02)
- All tasks are representative from the functional domain

3.1.7 Research (Alternative) Hypotheses

The following hypotheses support the research problem.

H_{a1}: Subjects answer more tasks correctly with the VITAMIN System than with the JIMI System. $\mu_{V1} > \mu_{J1}$

During the pilot studies, tasks of various complexities were identified. This research proposes that the number of tasks answered correctly for each type of task will be more with the VITAMIN system than with the JIMI System. These task-types are described in section 3.3.3, *Tasks*, below.

H_{a2}: Subjects have greater confidence in their correctly answered tasks with the VITAMIN System than with the JIMI System. $\mu_{V2} > \mu_{J2}$

H_{a3}: Subjects answer tasks correctly in less time with the VITAMIN System than with the JIMI System. $\mu_{V3} < \mu_{J3}$

H_{a4}: Subjects are more satisfied with the VITAMIN System than with the JIMI System. $\mu_{V4} > \mu_{J4}$

3.1.8 Testable Null Hypotheses

The following hypotheses are to be tested and possibly disproved in the analysis.

H_{01}: The mean number of tasks answered correctly with the VITAMIN system, μ_{V1}, will be less than or equal to the mean number of tasks answered correctly with the JIMI system, μ_{J1}. That is: $\mu_{V1} \leq \mu_{J1}$

Additional null hypotheses for the task-types are that the mean number of tasks answered correctly with the VITAMIN system, will be less than or equal to the mean number of tasks answered correctly with the JIMI system for each of the three task types.

H_{02}: The mean level of confidence in correctly answered tasks with the VITAMIN system will be less than or equal to the mean level of confidence in task answers with the JIMI system. $\mu_{V2} \leq \mu_{J2}$

H_{03}: The mean time to answer tasks correctly with the VITAMIN system will be greater than or equal to the mean time to answer tasks with the JIMI system. $\mu_{V3} \geq \mu_{J3}$

H_{04}: The mean level of satisfaction with the VITAMIN system will be less than or equal to the mean level of satisfaction with the JIMI system. $\mu_{V4} \leq \mu_{J4}$

3.2 Experimental Design

> In science one tries to tell people, in such a way as to be understood by everyone, something that no one ever knew before. But in poetry, it's the exact opposite. - Dirac [DIR77]

The research experiment is a within-subject randomly counter-balanced design (CBD) with 56 subjects. Each subject was tested on both the VITAMIN and JIMI systems. A within subject design was used because of anticipated high inter-subject variability with a small number of subjects. This design also allows a smaller number of total subjects to generate more case data. Varying the order of the treatments, administered to the subjects, counterbalances the experiment.

Approximately half of the subjects received the JIMI treatment first and the VITAMIN treatment second. A parallel set of tasks is used on the second interface to reduce the chance of performance improvement. The order is reversed for the other half of the users. The order of the treatments is randomly selected by JUICE.

The tasks for each treatment are parallel in the sense that two copies of the data sources are used for the experiment. The copies are identical except that the notional ship names have been changed. Similarly, the representative tasks are identical for both treatments except for the ship names. The experimental task lists for both treatments and the ship name mappings are contained in Appendix C.

3.2.1 Independent Variable

The single independent variable is the type of indirect manipulation interface system (IMIS). This research has two treatments or conditions. The control treatment is a traditional Java Indirect Manipulation Interface (JIMI) system to the agent-mediated legacy data. The experimental treatment is the VITAMIN system.

In the statistical analysis of the experimental results, the research goal is to determine whether the treatment, the VITAMIN system, makes a difference. The null hypothesis is that the independent variable, IMIS-type, does not make a difference. Therefore, the goal is to test the null hypothesis and claim statistically significant results.

3.2.2 Dependent Variables

The coherence of the interface systems is measured by the following dependent variables:

1. Number of correctly answered tasks (correctness)

2. User confidence in correctly answered tasks (confidence)

3. Time to answer tasks correctly (speed)

4. User satisfaction with the interface (satisfaction)

3.2.3 Pilot Study Experiments

Demonstration pilots of VITAMIN and JIMI prototypes were conducted to validate the instruments and adjust the methodology. Phase I of this research included formal demonstrations of the methodology and prototype systems at the Logistics Integration Agency and Mitre Corporation in the spring of 1998, and at West Point in the fall of 1998. Schafer, Rogers, and Marin [SCH98] describe the presentation at the International Multimedia Information Systems workshop (MIS'98) in Istanbul in the fall of 1998. Based on these prototypes, phases II and III included the two rounds of pilot study experiments described in this section.

The research design or structure is concisely described with the notation in Table 2. Elements of the design include the observations, treatments, groups, assignment, and time as described in Trochim [TRO99].

Symbol	**Description**
R	Subject groups randomly assigned to the order of treatments
O_D	Observation, Demographics
X_V	Treatment, VITAMIN
O_V	Observation, VITAMIN
X_J	Treatment, JIMI
O_J	Observation, JIMI
O_C	Observation, Comparison of VITAMIN and JIMI overall
O_E	Observation, Evaluation of JUICE

Table 2 Pilot Research Notation

The specific design used for the pilot study shown in Table 3 indicates the order of pilot experimental events from left to right. Each row corresponds to one observation case.

R	O_D	X_V	O_V	X_J	O_J	O_C	O_E
R	O_D	X_J	O_J	X_V	O_V	O_C	O_E

Table 3 Pilot Research Design Structure

During the pilot studies, the notes about the way subjects coped with the tasks and problems in conducting the experiment were recorded.

3.2.3.1 Representative Tasks

The tasks were taken from representative tasks provided by the US Army Logistics Integration Agency (LIA) AWR analysis group to find various information about readiness and availability of equipment in the afloat set of prepositioned stocks. These tasks were prepared according to given specifications from expert logisticians and analysts in the AWR domain. A copy of the experimental task list is at Appendix C. An example task is: "What is the Percent Fill of Pacing Items (ERC=P) on the ship named "Alexandria"?

These representative tasks were used in the experiment to address the first three research hypotheses: *correctness, confidence,* and *speed*. Eighteen tasks for each treatment were prepared and used for the pilot studies.

To mitigate threats to internal validity, the tasks are the same for each treatment except for the ship names. The independent treatment assumptions and verification tests are described in section 3.2.6 *Data Collection and Analysis*. Ship name mappings are listed in appendix C.

3.2.3.2 Subjects and Questionnaires

Phase II and Phase II included pilot studies conducted at West Point, NY in the spring and summer of 1999 with 63 subjects. All subjects were college students from the freshman psychology course. Their incentive for participating in the research included a possibility of extra credit as listed in Appendix F.

The experiments were administered to groups that varied in size from two to twelve subjects at a time. Thirty-two subjects in phase II were used to ensure the validity and reliability of the instruments. Thirty-one phase III subjects were used to further clarify the tasks and to gather input about possible biases according to perceived usefulness and perceived ease of use as advocated by Davis, Bagozzi, and Warshaw [DAV89] in the Technology Acceptance Model and also by Shneiderman [SHN97a] and Nielsen [NIE94].

For the pilot study, the subject background survey, at Appendix B, included nine questions to check the experience level of the subjects with computers in general and with data search and analysis tools in particular. According to the results of demographic surveys and their common academic core requirements, the subjects form a rather homogenous group in terms of computer experience. All of them use a computer nearly every day and have at least one of year experience in using computers.

The subject preference questionnaire to measure overall IMIS satisfaction included eight questions to evaluate each of the two interface systems, JIMI and VITAMIN. This overall user *satisfaction* instrument is found in Appendix D. The pilot study also presented the JUICE system evaluation survey at Appendix E.

Lack of time was the main problem with each of the pilot studies. Many subjects failed to complete the experiment in the allotted time. Numerous subjects spent the entire hour on one treatment and did not complete the rest of the experiment. Verbal instructions to "answer every other task question." were given during some of the pilot studies. These verbal instructions mitigated the time problem, but subjects were still rushed.

3.2.4 Dissertation Experiments

Based on the pilot results, the number of representative task questions was reduced from 18 to 12 for each treatment. Additionally, no demographic discriminators were identified in the pilot demographic surveys so the demographic survey, O_D, was reduced to two questions. Finally, JUICE proved to be a valuable and understandable enabler of the experiments and so the JUICE evaluation, O_E, was eliminated from the dissertation experiment as well. These changes to the dissertation experimental design allowed more time for the subjects to work on the research question.

Symbol	Description
R	Subject groups randomly assigned to the order of treatments
O_D	Observation, Demographics (2 questions)
X_V	Treatment, VITAMIN
O_V	Observation, VITAMIN
X_J	Treatment, JIMI
O_J	Observation, JIMI
O_C	Observation, Comparison of VITAMIN and JIMI overall

Table 4 Dissertation Research Notation

The dissertation experiment notation is shown in Table 4 and the design is shown in Table 5. Each row in Table 5 corresponds to one observation case used in the analysis.

R	O_D	X_V	O_V	X_J	O_J	O_C
R	O_D	X_J	O_J	X_V	O_V	O_C

Table 5 Dissertation Experiment Research Design Structure

Several supplementary changes were also incorporated into the dissertation experiment based on the pilot studies. These improvements include more descriptive background material, better instructions, and enhanced examples.

3.2.4.1 Representative Tasks

The number of representative tasks was reduced to twelve. The first three questions were example tasks with specific instructions to guide the subject to the correct answer. The nine remaining tasks were used for the analysis. Analysis of the pilot results in preparation for the dissertation proposal revealed that the subject's ability to answer the tasks seemed to correspond to the difficulty of the task. The enhanced IMIS, VITAMIN, appeared to offer greater advantages for harder tasks. Therefore, the dissertation research recognized several different types of tasks.

The dissertation research examined the two IMISs using three different types of tasks. The compound term "task-type" is used here to avoid confusion with the "type I and type II error" terminology employed in statistical hypothesis testing. As stated above, these task-types were established during the pilot experiments. They correspond to subject difficulty in performing representative tasks. The first type of task can be answered with one simple query. The second and third representative task-type questions require application and analysis of increasing complexity. The specific characteristics of each task-type are shown in Table 6.

Task-Type	**Characteristics**
Task type I	clearly specified, least detailed tasks identified by one query
Task type II	application of an additional query in one dimension required to solve the task
Task type III	most detailed, analysis of several queries required to answer the task

Table 6 Task-Type Characteristics

The division of tasks into types was based upon the levels of comprehension, application, and analysis required for answering each task. Bloom [BLO56] created a taxonomy for categorizing the level of abstraction required to answer tasks. A summary of Bloom's taxonomy for describing these competences and skills is shown in Figure 26.

Knowledge	Observation and recall of information
Comprehension	Understand and interpret meanings
Application	Use theories to solve problems
Analysis	Recognize patterns and components
Synthesis	Relate knowledge from several areas
Evaluation	Assess and verify theories

Figure 26 Bloom's taxonomy [BLO56]

Task-type I representative tasks have a straightforward and an accurate task definition that may be answered by observation from the high level aggregation presented by a single query. An example of such a Task-type I task is, "What is the authorized quantity of all items on the ship named "Key West"? Recognition and application of the proper detail dimension of the query is required to answer type II tasks. Task-type III tasks are more detailed and less clearly specified. They require the analysis of several query submissions that relate knowledge from several areas. A task-type III example is, "How many Infantry Rifle Companies are in the BNTF on the ship named "West Point"?

The dissertation experiment presented three example tasks followed by nine representative tasks. The three example tasks contained one task of each type. The first practice question was of task-type I, the second of task-type II, and the third of task-type III.

The order of the nine evaluated tasks was three task-type groups of three tasks each. The first group was of task-type I, the second of task-type II, and the third of task-type III. This order was used to provide the subjects with the simplest task-types at the beginning and then increase the complexity by task-type group. This order may result in improved performance with the later task-types, however it is impossible to answer type III tasks without knowing how to answer task-types I and II.

3.2.4.2 Subjects and Questionnaires

Fifty-six subjects participated in the phase IV dissertation experiments conducted in the fall of 1999 at West Point, NY. Fifteen subjects did not complete the exercises. Therefore, forty-one subjects were analyzed. The within-subject design with two treatments resulted in 82 experimental cases for analysis. The subject preference questionnaire includes eight questions to evaluate each of the two interface systems, JIMI and VITAMIN. This overall user *satisfaction* instrument is found in Appendix D. As with the pilot experiments, all subjects were college students from the freshman psychology course at West Point. Their incentive for participating in the research included a possibility of extra credit as listed in Appendix F. No attempt will be made to generalize the results beyond the homogenous demographics of the subject pool.

3.2.5 Research Procedures

The following research procedures, depicted in Figure 27, were used for the dissertation experiments. These same procedures were used for the pilot studies except where noted.

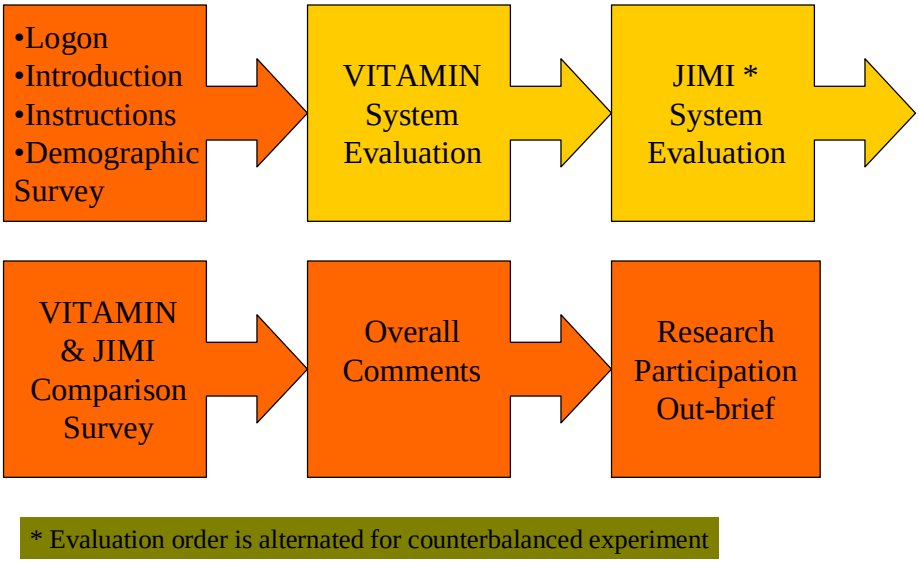


Figure 27 Java Usability Interface Comparison and Evaluation (JUICE)

The subjects sign a consent form and then connect to the experiment's web site. The web site contains the subject instructions. These instructions are repeated in Appendix A. The JUICE systems presents each subject with a background demographic survey. Then the JUICE system randomly assigns one of the treatments (VITAMIN or JIMI) and then the second treatment. During each treatment, the subject is presented with the interface system in the left display panel and the representative task on the right panel. The subjects use the interface system to answer the tasks and indicate their confidence in the answer. JUICE enables the recording of all subject's answers and task completion times for each task.

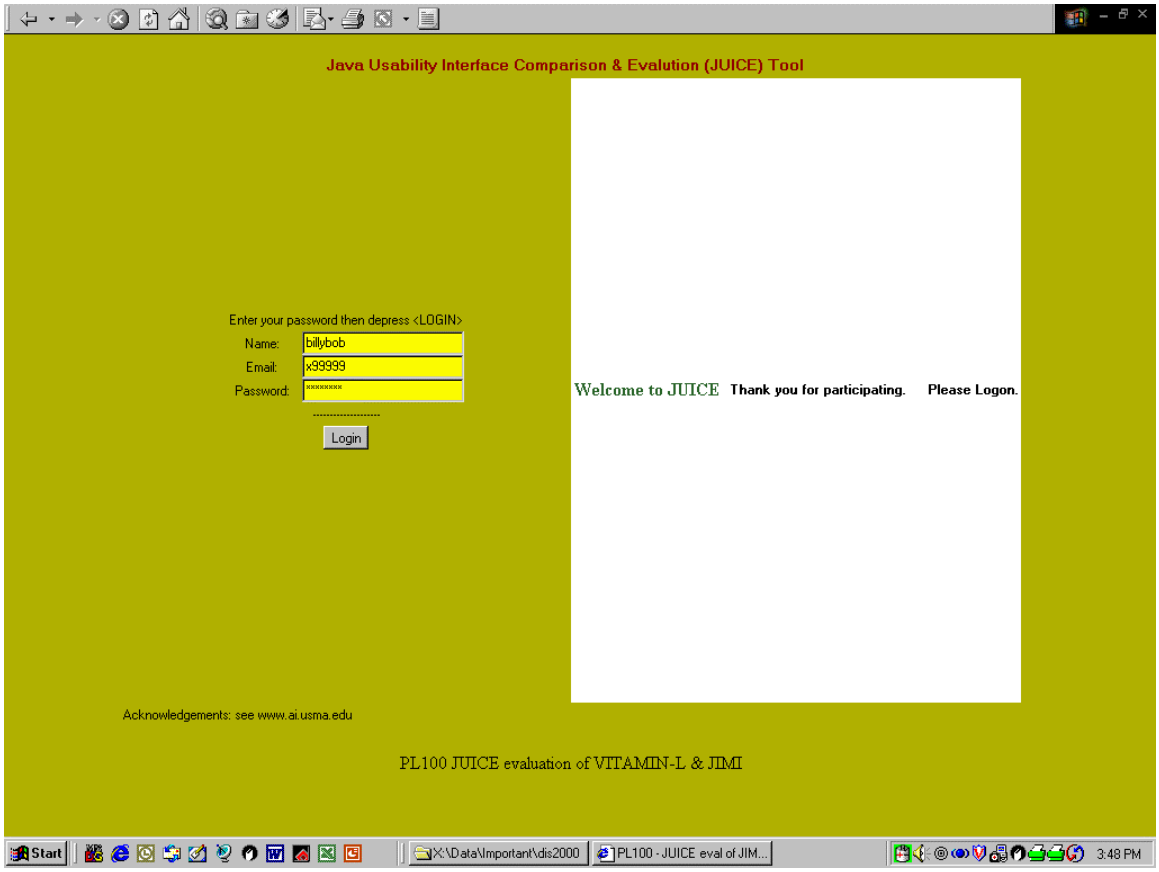


Figure 28 JUICE logon

As discussed above, 18 tasks were offered for each IMIS during the pilot studies. For the dissertation experiment, each subject was presented with twelve tasks. Three example tasks were provided at the beginning. These example tasks guided the subjects through using the IMIS for one question of each task-type. The remaining nine evaluated tasks for each interface consist of three groups of three questions for each task-type. Fifty minute are allotted for the experiment including the instructions and the questionnaires. The JUICE system guides the subjects through the experiment, collects subject actions, and collects timing and answers for later analysis. Figures 28 – 36 provide screen shots with explanations of a sample JUICE experiment session.

JUICE initial logon screen linked directly from the instructions web page. A copy of these user instructions is at Appendix A.

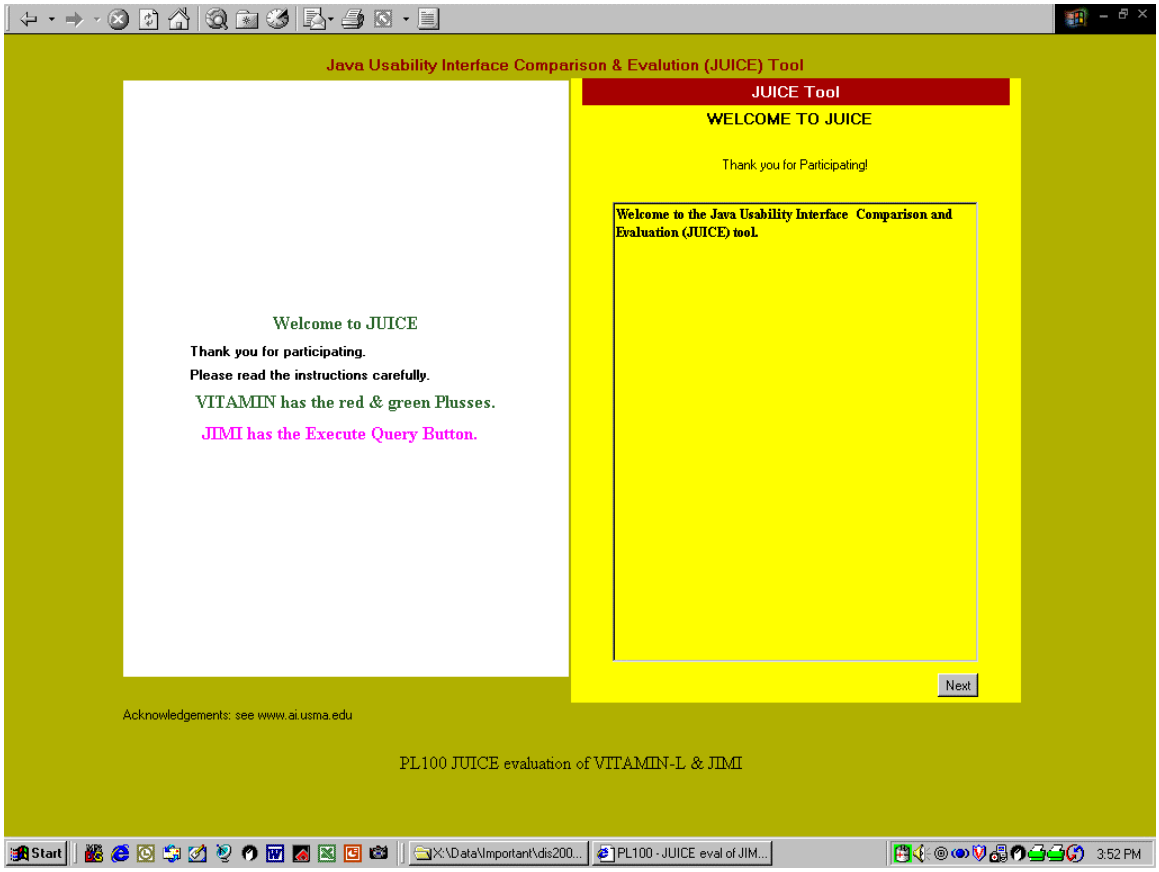


Figure 29 JUICE Welcome

After logging in, the subject is presented with the welcome screen. The subjects are not required to log in separately to JIMI and VITAMIN; JUICE logs them in using the information they initially provided.

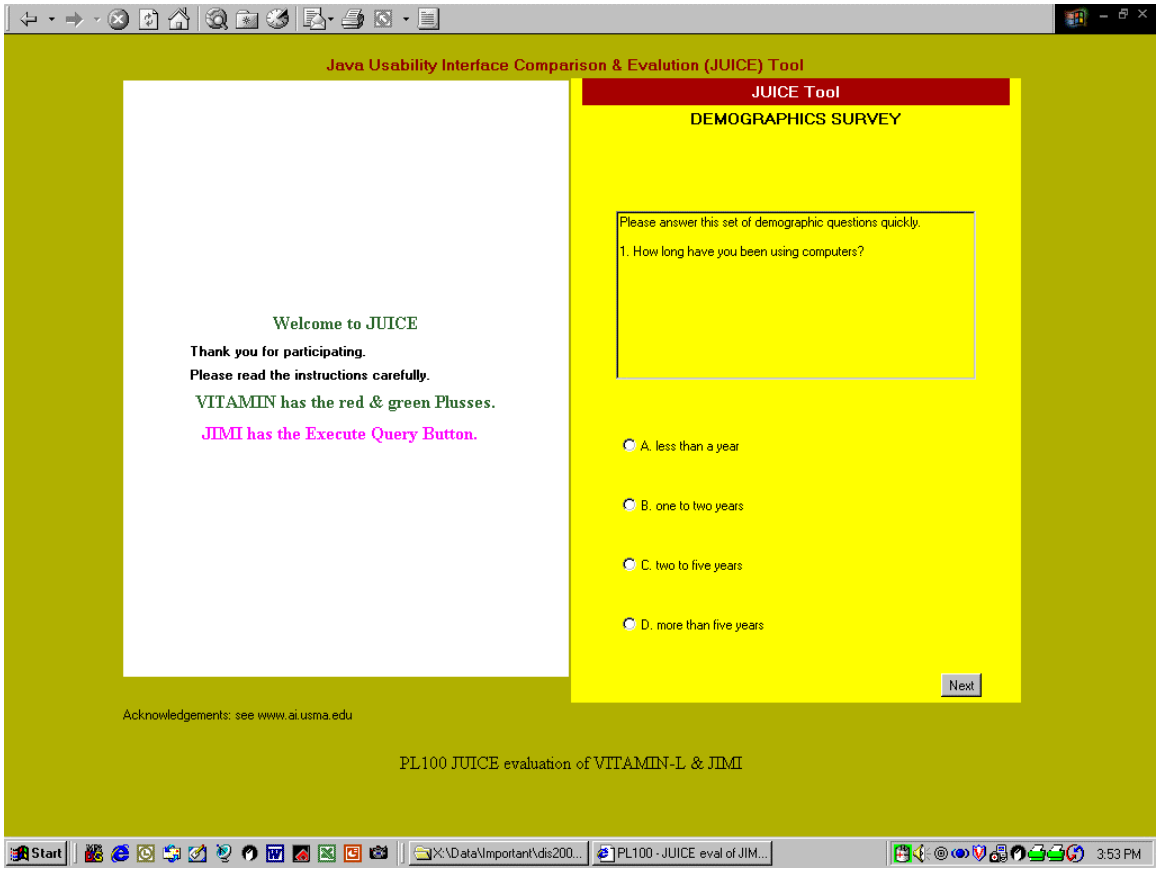


Figure 30 JUICE Demographic Survey

The first experimental observation, O_D, is the background demographics survey. JUICE presents this survey and requires that the subjects select one of the radio-buttons (A. –D.). The subject then selects the "Next" button to continue. If the subject tries to select the "Next" button without answering the question, JUICE displays a message reminding them to select an answer.

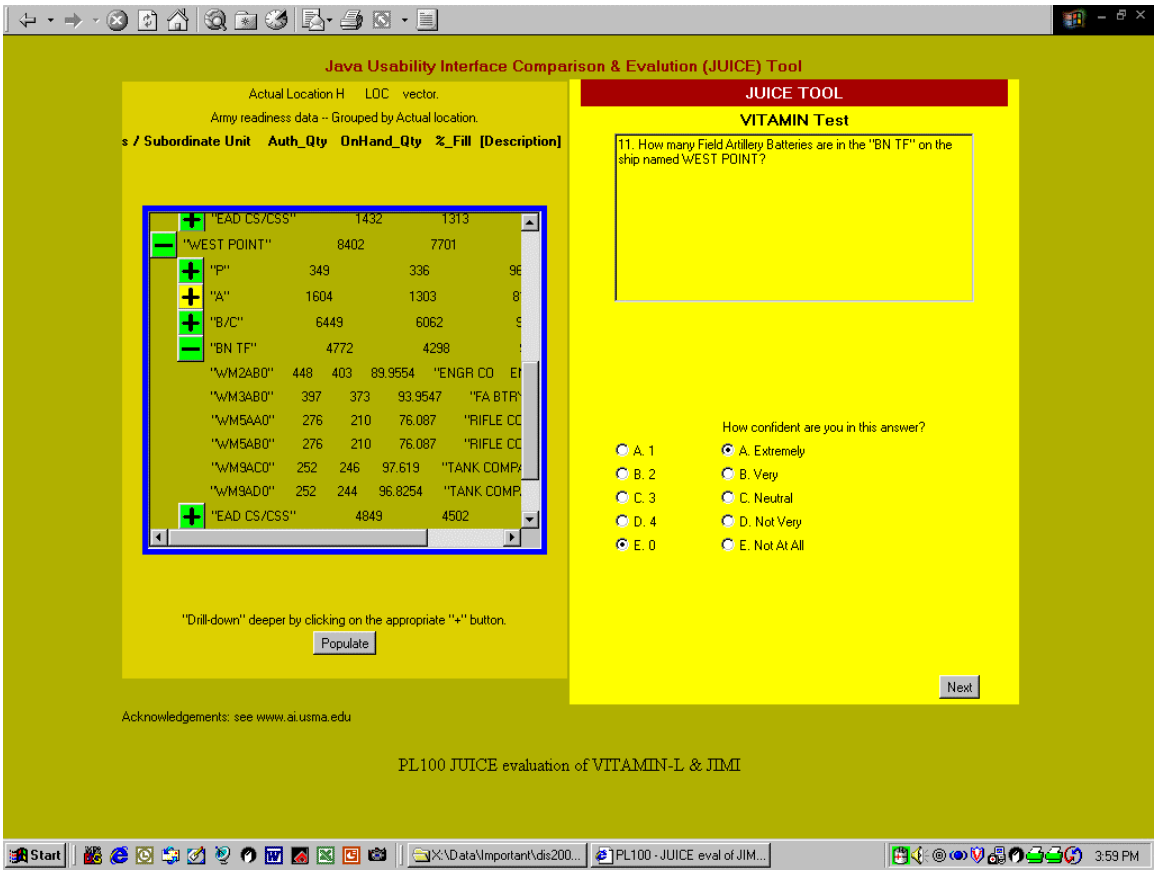


Figure 31 JUICE Task with VITAMIN

Following the background survey, JUICE randomly assigns the first treatment – either VITAMIN or JIMI and presents the subject with representative tasks. VITAMIN is the IMIS treatment in the example above. As mentioned earlier, the first three tasks are examples that provide systematic instructions for using the specific IMIS to answer the task question.

Twelve tasks are provided for each IMIS. Following the three example tasks, JUICE presents three type I tasks, then three type II tasks, and finally three type III tasks. This order provides representative tasks of increasing complexity in terms of number of queries required. In addition to requiring the subjects to generate more queries with the IMIS, this order demands increasing analysis to answer the representative tasks.

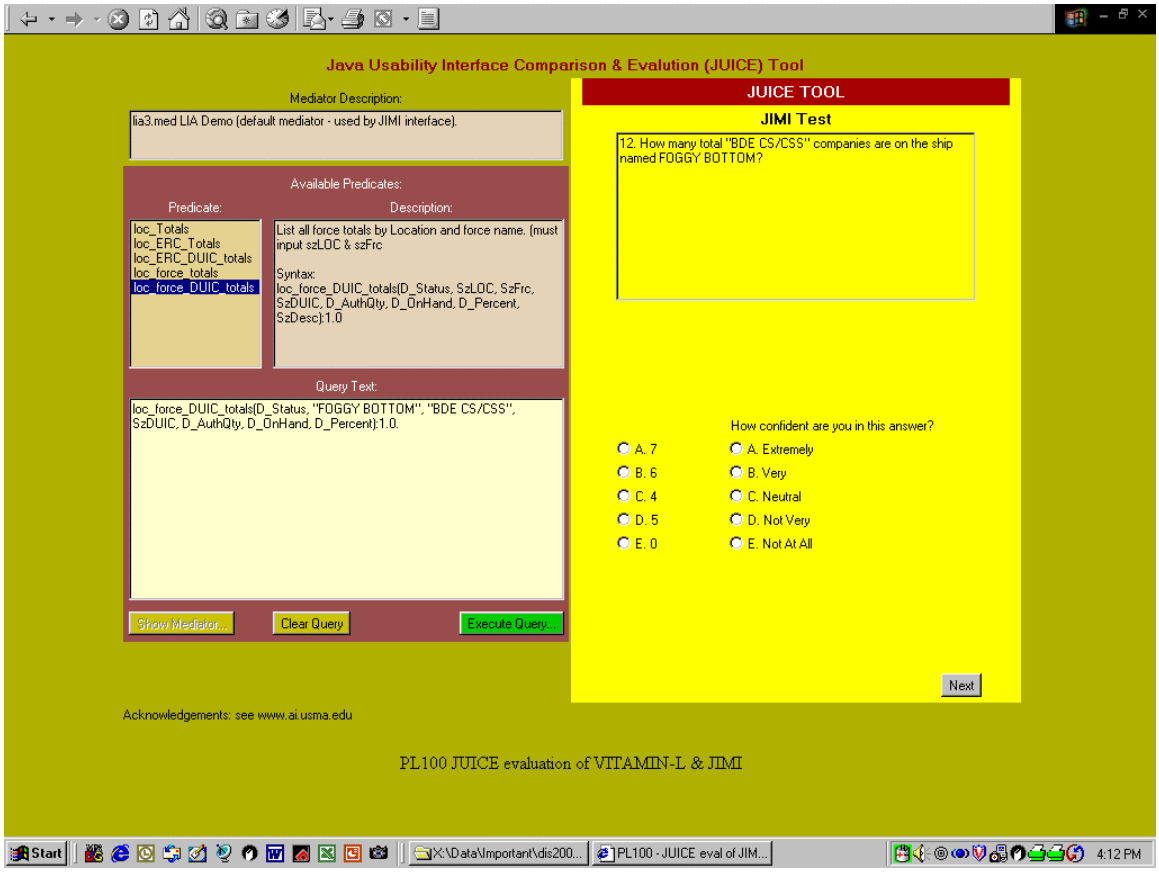


Figure 32 JUICE Task with JIMI

After the subject completes the first treatment, JUICE assigns the second treatment, using the other IMIS, with a parallel set of representative tasks. For each treatment, the IMIS is presented in the left panel and the task questions are presented in the right panel.

For each task, the subjects use the IMIS on the left, JIMI in the example above, to find the correct answer to the representative task that JUICE presents on the left half of the screen. As with the demographic survey, the subjects must select an answer and a confidence value in order to continue to the next question by selecting the "Next" button.

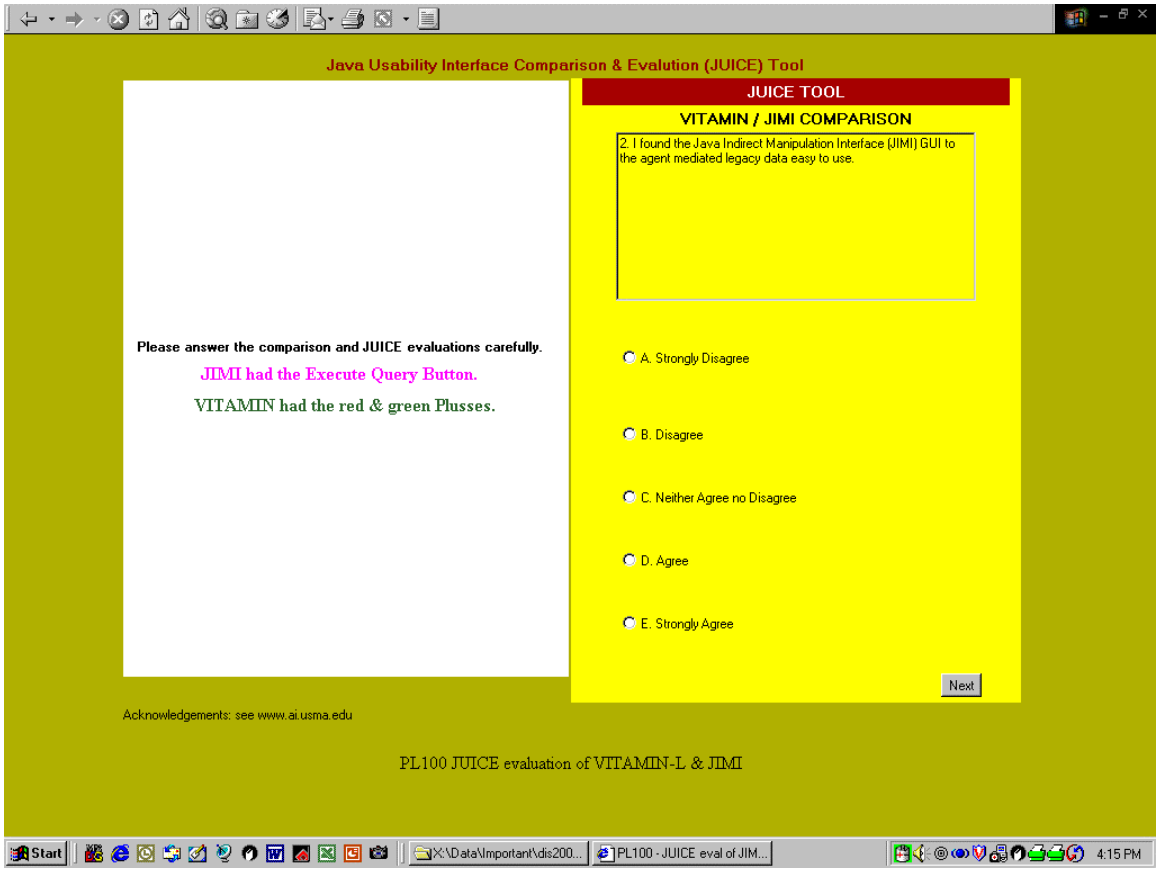


Figure 33 VITAMIN / JIMI Comparison

Following both treatments, the comparison instrument is presented. Again, the subjects must provide a response in order to continue. The pilot studies also revealed that several subjects could not recall which IMIS was VITAMIN and which was JIMI. The panel on the left displays a reminder.

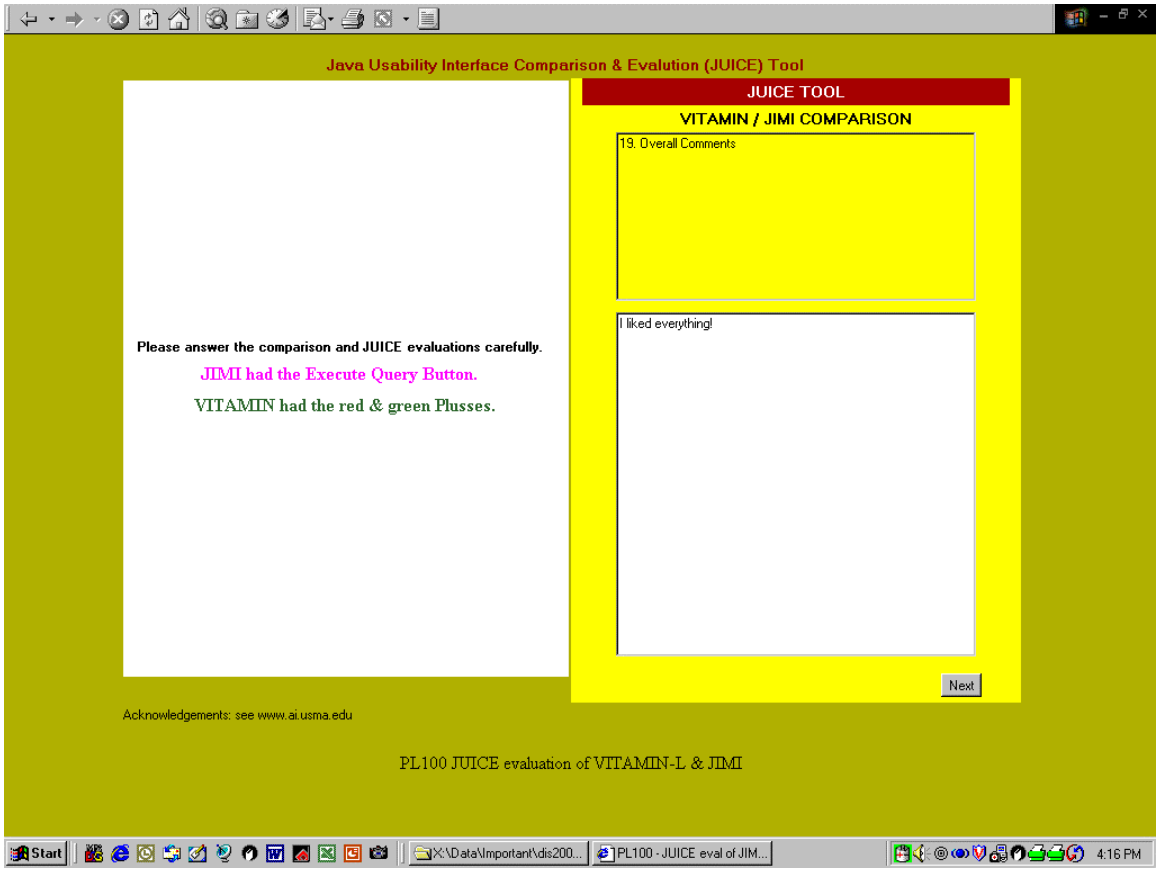


Figure 34 JUICE VITAMIN JIMI Comments

The last three questions offer the subjects the ability to enter free text responses about JIMI, VITAMIN, and the overall system.

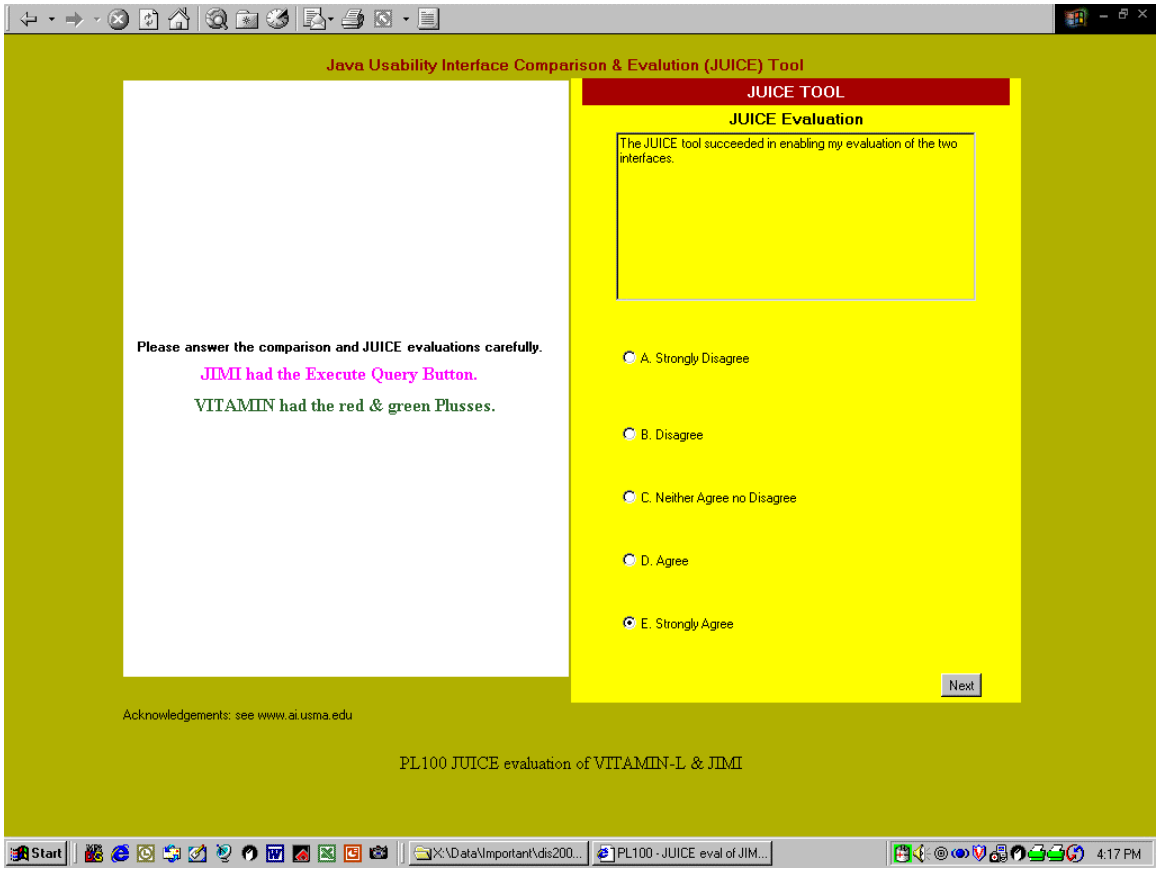


Figure 35 JUICE Evaluation

Finally, during several of the pilot trials, a formative evaluation of JUICE was performed. The JUICE evaluation was not conducted for the dissertation experiments. Analysis of the pilot evaluations resulted in improvements to JUICE and the experimental procedures.

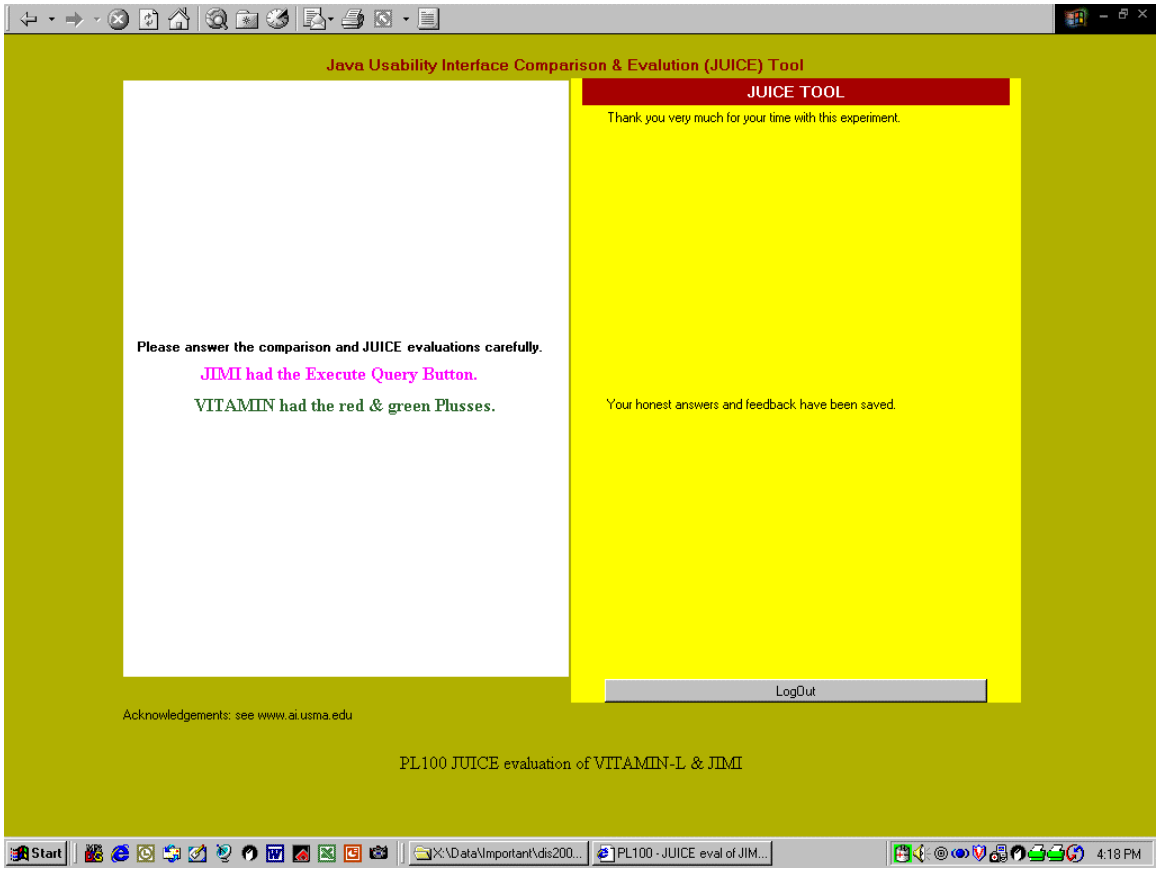


Figure 36 JUICE Logout

The final screen is shown above.

The research procedures were refined based on demonstrations and pilot studies. In addition to procedural improvements, the research was strengthened based on input from the functional domain community and the subject population.

3.2.6 Data Collection and Analysis

The JUICE system ensures that all data is collected and stored on the server through the HERMES mediator agent. Specifically, answers to all questions, user actions, timing, and identification information are collected by JUICE. The timing information is used to calculate the total user time required to complete each representative task. The system time delays are subtracted from the total time for each query.

The results are analyzed with descriptive statistics to include the mean, standard deviation, minimum, and maximum. These results are presented in box and whisker plots.

For within-subjects designs, a key advantage is that the number of cases is the product of the number of subjects and the number of treatments. For a single independent variable with two treatments, the number of research cases is equal to twice the number of subjects. The key assumption in this design is that each treatment is independent. However, the order of treatment may affect the results. This is called an order effect or carry-over effect. It may result in a practice effect whereby the subjects do much better with the second treatment or a fatigue effect whereby the subjects do much worse on the second treatment.

This order effect is often referred to as the "washout assumption" in clinical drug trials because the assumption is that the drug treatment has completely "washed-out" of the subject. Randomly counterbalancing the design mitigates this risk. Randomly counterbalancing the order of the single independent variable results in complete counterbalancing for this research. This results in a strong mitigation of the order effect and assures random observations from the pool of subjects. The most conservative test of this washout assumption is to discard all but the first treatments and examine the degree to which the treatment effect is still valid. This washout assumption test is performed for each hypothesis.

The hypotheses are tested with the independent samples t-test since information is not available about the population. As discussed above and in Appendix F, the subject selection is not truly random; however, the subjects are selected in accordance with accepted human research and usability evaluation practices. Demographic differences between the subjects and the actual analyst population should not be problematic for the purposes of testing the stated hypotheses. Additionally, according to conventional usability evaluation procedures, the threat to external validity or the generalizability of the results imposed by the homogenous subject pool will not critically impair the applicability of the analysis.

For the hypothesis tests, each observation case consists of the one randomly assigned treatment and the corresponding observations. This corresponds to a single row in the experimental structure discussed above. The t-test assumes that the observation cases are independent random samples from normal distributions. According to the Central Limit Theorem, means of samples of size n>= 30 will be approximately normally distributed. According to Sternstein [STE94], the t-distribution is nearly identical to the normal distribution for n>=30 and is more conservative for a small sample sizes. The t-test is also robust to departures from normality. In this research, the t-distribution is also selected in order to compare directly the entire set of cases (where n>30) with the smaller sample washout assumption test (where n < 30).

Statistical techniques are also used on the survey and confidence data although these data are categorical or ordinal rather than quantitative interval variables. According to Jaccard and Wan [JAC96], "For many statistical tests, rather severe departures (from intervalness) do not seem to affect Type I and Type II errors dramatically." Garson [GAR99] says: "Use of ordinal variables such as 5-point Likert scales with interval techniques is the norm in contemporary social science." Common usability instruments, including QUIS and SUMI, make use of this type of analysis.

As discussed above, for the analysis of the survey data, the Likert scales' ordinal values are treated as interval. The task confidence questions scale from "A. Extremely" to "E. Not at All." These are mapped from the interval 1.0 to 0.0 at 0.25 steps. The survey questions scale from "A. Strongly Disagree" to "E. Strongly Agree." These are mapped from the interval 0.0 to 1.0 at 0.25 increments. These mappings are presented in Appendices C and D, respectively.

Finally, according to Patton [PAT90], statistical portrayals must be interpreted and given meaning. To assist this analysis, open-ended survey questions were also provided to add qualitative richness to the quantitative data. These responses capture the subjects' responses in their own terms and allow a deeper view of the context, which is an important element of usability inquiry. The qualitative data is analyzed for common

themes and relationships to the quantitative data. Correlation effects are not addressed in this research and no correlation-effect implication is made for the hypotheses.

3.2.7 Reliability and Validity

According to Kirakowski [KIR99], the ability of the questionnaire to give the same results when filled out by concurring subjects in similar circumstances is known as reliability. A measure is reliable if the application of the measure yields reproducible results. The degree to which the questionnaire is actually measuring data that it was designed to collect, and not being misinterpreted by subjects is known as validity.

All representative tasks were obtained from, and reviewed by, functional domain experts to assure content validity. The questionnaires used in this research were based on similar validated instruments including those investigated by Chelimsky [CHE91]; Chin, Diehl, and Norman [CHI87]; [CHI88]; Kirakowski and Corbett [KIR93]; Nielsen [NIE94]; Porteous, Kirakowski, and Corbett [POR93]; Shneiderman [SHN97a].

Task and survey questions were refined and validated by the subject responses to the systematic pilot studies so that all subjects had the same interpretation to assure instrument reliability. The questionnaires and task lists were also revised according to the recommendations and problems encountered during the pilot studies. Reliability and validity analyses of the survey instrument are included in Chapter 4.

3.2.8 Materials

The experimental materials include:

- A networked workstation with a web browser to access the instructions and to run JUICE
- Agent mediated Heterogeneous AWR Legacy Data Sources (HALDS)
- A subject background survey
- A set of tasks to be performed by the subjects for each interface system

- A traditional query interface system (JIMI)
- A VITAMIN system interface
- A subject preference questionnaire
- A JUICE system evaluation survey

3.3 Ethical Considerations and Issues

Ethics is extremely important in computer science according to Granger et al. [GRA97]; Huff and Martin [HUF95]; Martin [MAR97]; Reddy [RED95] and Routio [ROU99a]. All research for this dissertation was conducted in accordance with the GWU Code of Academic Integrity. Human research was conducted in accordance with USMA policies. Appendices F and G contain the detailed USMA policies for the use of human research subjects.

The author is also aware of the ethical anomaly in many scientific studies concerning possible consequences of this research. While it is important to consider the ethical considerations of research on the subjects, the standard "informed consent procedures" are completely uncontroversial for an experiment that has no chance of harming anyone. What may be controversial is the possible consequences of the research if a foreseeable outcome may make it easier to "make people suffer."

Professor Nagy, a member of the Supervisory Committee, posed the following questions: If any country can wage war with near-zero casualties, are they more likely to do so? If research makes it more efficient, does it raise the prospect that more people will suffer? People likely to suffer include friendly and enemy soldiers and non-combatants. A survey of historical and moral perspectives on these questions is offered in Appendix M.

In short, national policy and the consequences of public and international opinion have greater impact on decisions to use force than the ease with which it may be applied. Technological advances such as precision weapons probably reduce suffering rather than increase it. Finally, the scientist who conducts research within moral parameters cannot

prevent the redirection of the work and no prophylactic measure can prevent future mischief.

Professor Nagy commends the author for thoroughly addressing ethical considerations, however, his dissent with the analysis is at www.gwu.edu/~nagy/jscope2001.htm.

3.4 Methodology Summary

Chapter Three presented explanations and rationale for the three systems developed for this research, JIMI, VITAMIN, and JUICE. Specific implementation and development details were described. The research problem was reviewed and the research and null hypotheses were introduced. In addition, the experimental design, to include the variables, tasks, procedures, and analysis were discussed in detail. Finally, the ethical issues concerning human subject research and the possible consequences of scientific research were considered. The next chapter tabulates the experimental results and presents the statistical analysis of these results.

Chapter 4: RESULTS AND ANALYSIS

> Data is what distinguishes the dilettante from the artist - Higgins [HIG88].

Chapter Three provided descriptions of the research systems, implementation details, and the experimental methodology. These research systems, based on the theory in chapter two, were used to conduct a set of experiments to test this theory. The results of these experiments and the statistical analysis are presented in this chapter.

Fifty-six subjects took part in the dissertation experiments; however, fifteen did not complete the exercises. Therefore, forty-one subjects were analyzed. The within-subject design with two treatments resulted in 82 experimental cases for analysis. All of the quantitative data used for this analysis is at Appendix J; the qualitative data is at Appendix K.

The research questions and the empirical measurement of the concept of coherence are reviewed. The analysis procedures are explained. Each of the first three coherence hypotheses, *correctness*, *confidence*, *and speed*, is then examined in detail in subsequent sections. This analysis includes presentation of the descriptive statistics, hypothesis test, and washout assumption test. The results are also presented by task-type in this chapter for these hypotheses. Additionally, the same analysis procedure is reported for each representative task question in Appendix L. Washout assumption tests by question and task-type are also reported in Appendix L. The results of these three additional analyses, (1) washout by task-type, (2) by question, and (3) washout by question are briefly summarized in their respective sections.

The fourth hypothesis, *satisfaction*, is then examined with an analysis of the descriptive statistics, hypothesis test, washout assumption test, analysis of the individual satisfaction measures, and a reliability analysis.

Following the quantitative analyses, the qualitative results are reviewed for themes to add richness from the subjects' own characterizations. Moreover, the themes that emerge from this open-ended survey data are analyzed for possible support of the statistical portrayals. Finally, a summary of the analysis is presented in the last section.

4.1 Research Questions

Actual subject questions for all parts of the study are in the appendices listed in Table 7.

Appendix	**Questions**
Appendix A	Subject Instructions
Appendix B	Subject Background Survey
Appendix C	Experimental Task Lists
Appendix D	User Preference Survey
Appendix E	JUICE system Evaluation Survey

Table 7 Research Questions

The measures of coherence defined for this research correspond to representative tasks performed *correctly*, *confidently*, *speedily*, and overall system *satisfaction*. The empirical measures used for this analysis are computed for each of these factors as follows:

Correctness corresponds to the number of all tasks the subject answered correctly. The results and analysis for this correctness hypothesis is reported in section 4.3.

Confidence corresponds to the confidence rating the subject gave for each correctly answered question divided by the number of correctly answered questions. These results are reported in section 4.4.

Speed corresponds to the average time the subject spent answering all questions correctly. This is calculated by subtracting the system query time, which averages about 10 seconds per query from the time spent responding to the representative task. The time spent on

each correctly answered question is summed and the total is divided by the number of correctly answered questions. These results are reported in section 4.5.

Satisfaction corresponds to the subject's overall satisfaction rating for the system under consideration. This total satisfaction score is the sum of the scores for each attribute. The score is also provided for each of the eight satisfaction attributes. These results are reported in section 4.6. The reliability analysis of the survey instrument is reported in this section as well.

Additional results are reported in the first three sections mentioned above. These results include the analysis by task-types I, II, and III, and the washout assumption test for the hypothesis. Further outcomes for each hypothesis are reported in Appendix L. These outcomes include: the washout assumption test by task-type, the analysis of each representative task question individually, and the washout assumption test for each question individually. The *satisfaction* hypothesis relates to an overall rating of the IMIS, therefore, no data is available by question or task-type, and there are no washout assumptions.

4.2 Quantitative Analysis Procedures

For each hypothesis, summary descriptive statistics are reported along with a box-plot. The descriptive statistics include number of cases, minimum, maximum, mean, and standard deviation. The summary box-plot, or box-and-whisker-diagram, is based on the median, quartiles, and extreme values. The box represents the interquartile range that contains the central 50% of values. In some cases, the scale of the diagram prevents the display of the interquartile range for both treatments. There are lines, called whiskers, which extend from the box to the highest and lowest values, excluding the outliers. A dark line across the box indicates the median as described in SPSS [SPS99].

This box-plot display is useful to check that the median is approximately in the center of each box. It also displays whether the spreads of the two groups are similar and whether there are outliers. Outliers may serve to move the means closer together than the medians, and their presence might hide true differences. The reverse may also be true –

the presence of a few outliers may cause the means to differ more than the medians, which makes the test results appear more significant than the bulk of the data indicate.

Following the summary descriptive statistics and box-plot, the hypothesis test is calculated. The experimental design is a two-group, posttest-only random experimental design. The independent variable is the type of indirect manipulation interface system (IMIS). The independent variable divides the groups. For the purpose of the analysis, each dependent variable is treated separately. The treatment for each dependent variable is measured as two distributions, each with an average and a variation. The effect of the treatment, the t-value, is the difference between the means of two distributions divided by the variability around the means. This variability is called the standard error of the difference.

There are three different ways to estimate the treatment effect of this type: independent t-test, one-way ANOVA, and regression analysis. Regression analysis is the most general. It regresses the posttest values onto the dummy-coded treatment variable (Z). The resulting t-value for the slope coefficient is the same number resulting from the independent t-test. The square of the t-value is equal to the F-value that results from the one-way ANOVA. ANOVA is used to extend the comparison of means to more than two samples. These approaches evolved independently. They have been shown to be mathematically equivalent according to Trochim [TRO99]. This research uses the independent t-test since regression and ANOVA extensions are not required. The goal of the analysis is to determine whether the effect is statistically significant. The independent samples t-test supports this determination.

The level of risk or *alpha* level for all hypotheses is set to 0.05. This level is commonly specified in usability research. This means that five times out of one hundred a statistically significant difference between the means is found by chance when there actually is none. This mistaken failure to reject a false claim is called a Type II error. (A Type I error is when a true claim is mistakenly rejected.)

The standard independent samples t-test, called a *pooled-variance* t-test, assumes equal variance in the populations. The box-plot can give an indication as to the validity of this assumption. If the larger variance of the two samples is no more than three times the smaller variance, then it is still legitimate to use the t-test according to Clark-Carter [CLA97]. If the variance is greater than three times the smaller variance, then the Smith-Satterthwaite procedure is used as described in Arnold and Milton [ARN95]. To verify this assumption, Levene's test for equality of variance is computed. If the observed significance level for this test is less than 0.05, then the *separate variance* t-test is used. The penalty for violating the equal variance assumption is diminished degrees of freedom, which means that the effective sample size decreases as shown in SPSS [SPS99]. This reduced degrees of freedom results from the requirement to abandon the *pooled variance* t-test in favor of the Smith-Satterthwaite procedure used in the *separate variance* t-test in which equal variance is not assumed.

Based on the results of the test for equality of variance, either the *pooled-variance* or *separate variance* t-test is computed. Each hypothesis has directionality; therefore, the one-tailed probability values are used. The one-tailed probabilities are obtained by dividing the standard two-tailed t-probabilities by two. The results of the hypothesis test is reported along with a statement of its meaning. Correlation effects are not addressed in this analysis.

The same analysis procedure (descriptive statistics, box-plot, t-test) is applied for each task-type and for each question of the first three hypotheses. The most conservative test of the washout assumption is to discard the data from the second treatment for each subject and examine the degree to which the treatment effect is still valid. This analysis with the descriptive statistics, box-plots, and t-tests are provided for the test of the washout assumption for each hypothesis.

Then remaining sections of this chapter present the results and analysis of each hypothesis followed by the qualitative results and analysis summary.

4.3 Number Of Correctly Answered Tasks

The *correctness* factor corresponds to the number of all tasks that the subject answered correctly for the given treatment.

4.3.1 Descriptive Statistics - Correctness

The maximum number of tasks that could be answered correctly is nine, excluding the first three sample tasks.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	5	9	7.8	1.4
VITAMIN	42	5	9	8.6	0.8

Table 8 Number of Correctly Answered Tasks

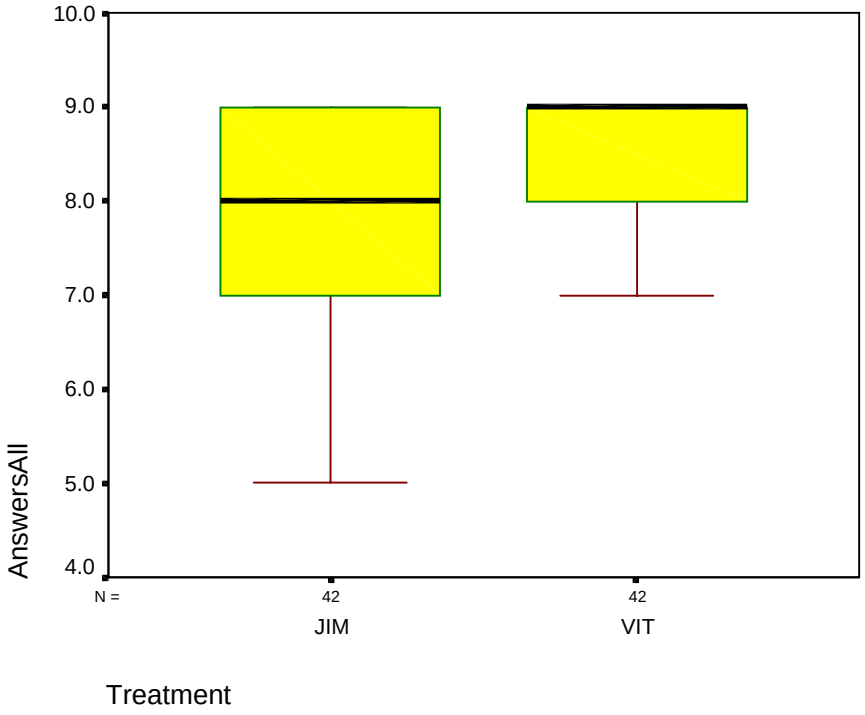


Figure 37 Number of Correctly Answered Tasks Box-plot

Examination of Table 8 and visual inspection of Figure 37 indicate that there is a difference between the two means and that the variance does not appear similar.

4.3.2 Hypothesis Test - Correctness

The observed significance level for Levene's test for equality of variance is less than 0.05 (L sig <<.001) so the *separate variance* t-test is used. The calculations for Levene's test are reported in Appendix L. The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean number of correctly answered questions (t=-3.0, df=64, p<=.002). Therefore, the null hypothesis is rejected and the first research hypothesis is accepted:

H_{a1}: Subjects answer more tasks correctly with the VITAMIN System than with the JIMI System.

Three additional null hypotheses are that, for each of the three task types, the mean number of tasks answered correctly with the VITAMIN system is less than or equal to the mean number of tasks answered correctly with the JIMI system.

4.3.3 Type I Tasks - Correctness

Following the three example questions, the first three tasks (questions 4, 5,and 6) were the simplest tasks of task-type I. The descriptive statistics, box-plot, and hypothesis test are presented below for this task-type.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	1	3	2.7	.60
VITAMIN	42	1	3	2.8	.51

Table 9 Number of Correctly Answered Tasks – Type I

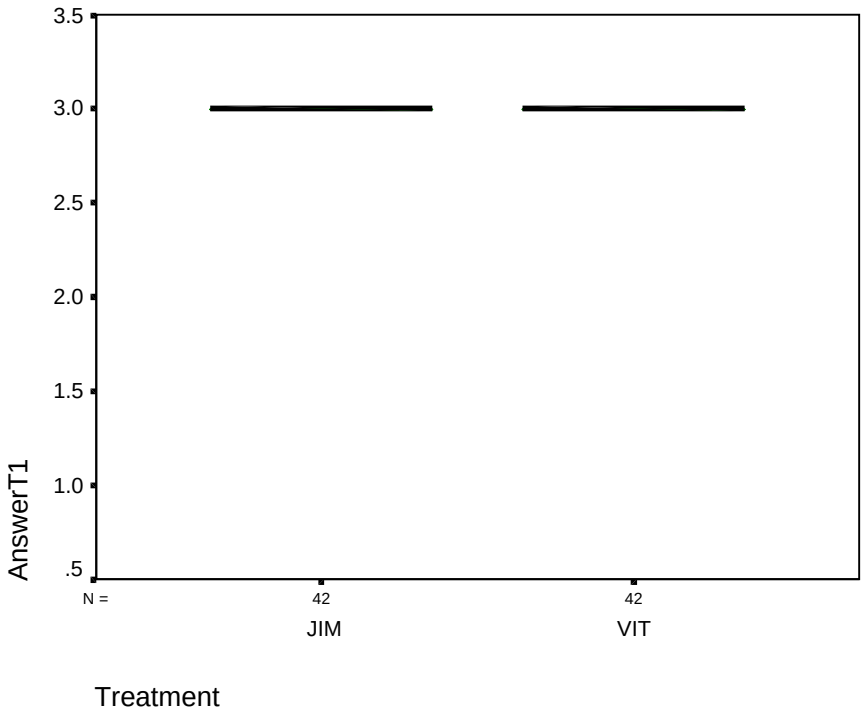


Figure 38 Number of Correctly Answered Tasks – Type I Box-plot

Examination of Table 9 and visual inspection of Figure 38 indicate almost no difference between the two means and that the variance appears similar; (L sig =.15). The one-tailed independent samples *pooled variance* t-test found no statistically significant differences in the mean number of correctly answered questions (t=-0.79, df=82, p<=.21). Therefore, the analysis fails to reject the null hypothesis for task-type I correctness.

4.3.4 Type II Tasks - Correctness

The second three tasks (questions 7, 8, and 9) were of task-type II. The descriptive statistics, box-plot, and hypothesis test are presented below.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	1	3	2.7	.55
VITAMIN	42	1	3	2.9	.40

Table 10 Number of Correctly Answered Tasks – Type II

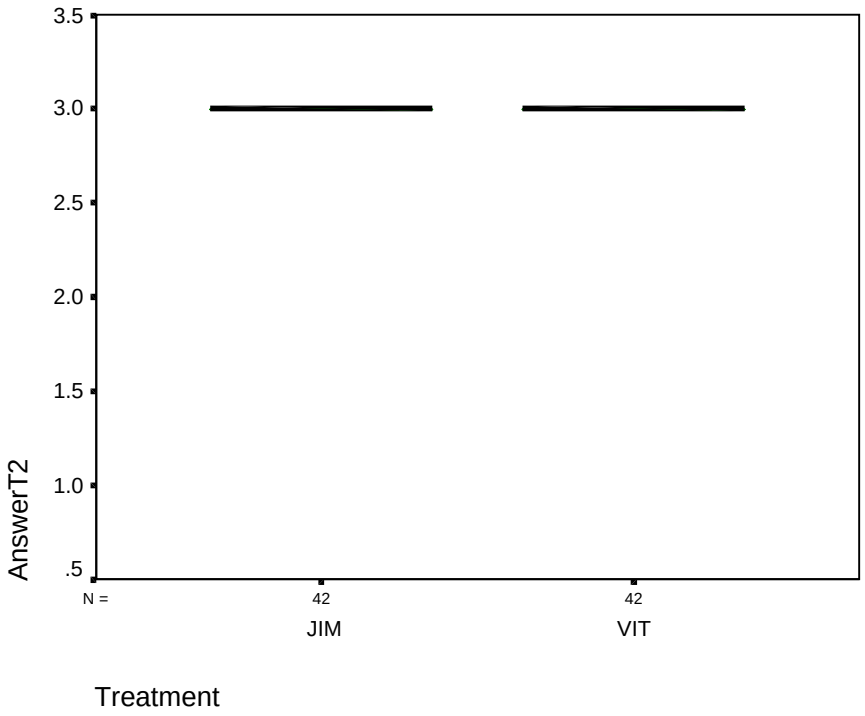


Figure 39 Number of Correctly Answered Tasks – Type II Box-plot

Examination of Table 10 and visual inspection of Figure 39 indicate almost no difference between the two means and that the variance appears similar; however due to outliers and extreme values L sig =.003. The *separate variance* t-test found no statistically significant differences in the mean number of correctly answered questions (t=-1.59, df=74, p<=.06). The analysis fails to reject the null hypothesis for task-type II correctness.

4.3.5 Type III Tasks - Correctness

The third and last three tasks (questions 10, 11, and 12) were of task-type III. The descriptive statistics, box-plot, and hypothesis test are presented below.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	1	3	2.4	.80
VITAMIN	42	2	3	2.9	.30

Table 11 Number of Correctly Answered Tasks – Type III

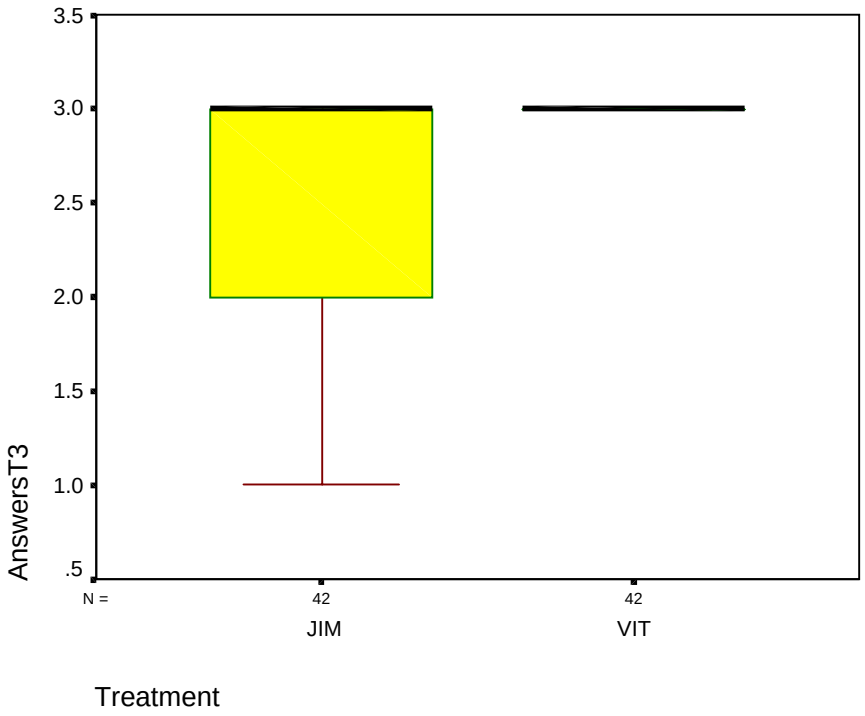


Figure 40 Number of Correctly Answered Tasks – Type III Box-plot

Examination of Table 11 and visual inspection of Figure 40 indicate almost no difference between the two means but the variance does not appear similar; (L sig <<.001). The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean number of correctly answered questions (t=-3.61, df=52, p<<.001). Therefore, the null hypothesis is rejected or task-type III correctness.

Overall, these three additional task hypotheses lend support to the advantage of the VITAMIN system over the JIMI system. There was almost no difference between the number of correct answers with the simplest Type I tasks (p<=.21). The Type II tasks showed differences that were not statistically significant (p<=.06 at alpha = .05), and the Type III tasks showed statistically significant differences (p<<.001).

4.3.6 Washout Assumption Test - Correctness

Discarding the second treatment results in 42 cases, 23 when JIMI was the first treatment and 19 when VITAMIN was the first treatment, as shown in Table 12.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	23	5	9	7.7	1.4
VITAMIN	19	8	9	8.7	.48

Table 12 Number of Correctly Answered Tasks – Washout

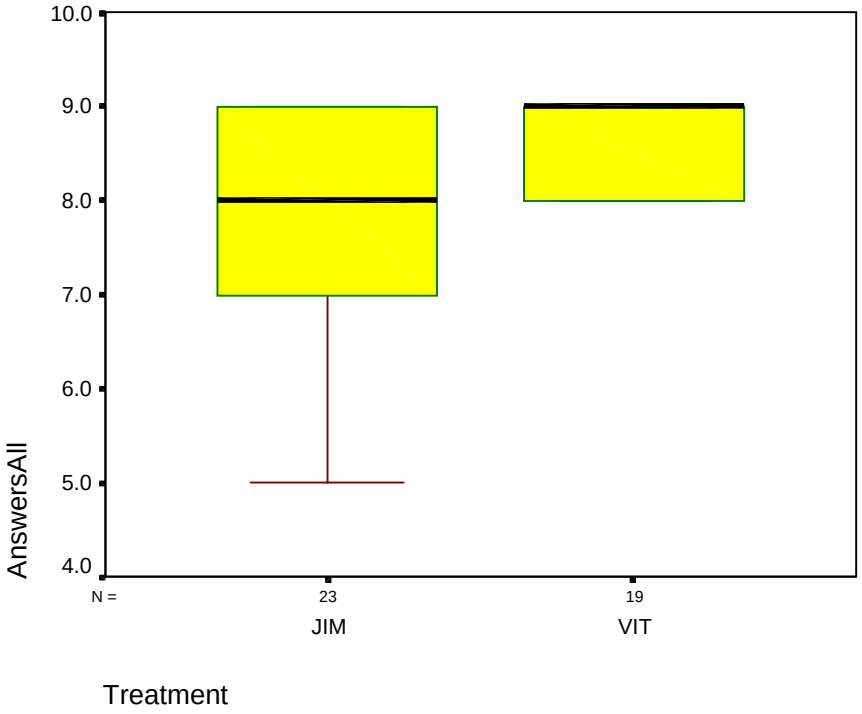


Figure 41 Number of Correctly Answered Tasks – Washout Box-plot

Examination of Table 12 and visual inspection of Figure 41 indicate that there is a difference between the two means and that the variance does not appear similar; (L sig =.001). The *separate variance* t-test found statistically significant differences in the

mean number of correctly answered questions (t=-2.99, df=28, p<=.003). Therefore, the null hypothesis is rejected and the washout assumption is valid for the first hypothesis.

4.4 User Confidence In Correctly Answered Tasks

The *confidence* factor corresponds to the confidence rating the subject gave for each correctly answered question divided by the number of correctly answered questions. The confidence rating scale is shown in Table 13.

How confident are you in this answer?	Analysis Value
A. Extremely	1.00
B. Very	0.75
C. Neutral	0.50
D. Not Very	0.25
E. Not At All	0.00

Table 13 Confidence Analysis Values

4.4.1 Descriptive Statistics – Confidence

Accordingly, a value of 1.0 indicates "Extremely" and a value of 0.0 indicates "Not at all" confident in the descriptive results shown in Table 14.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	.33	1.00	.87	.18
VITAMIN	42	.50	1.00	.94	.12

Table 14 Confidence in Correctly Answered Tasks

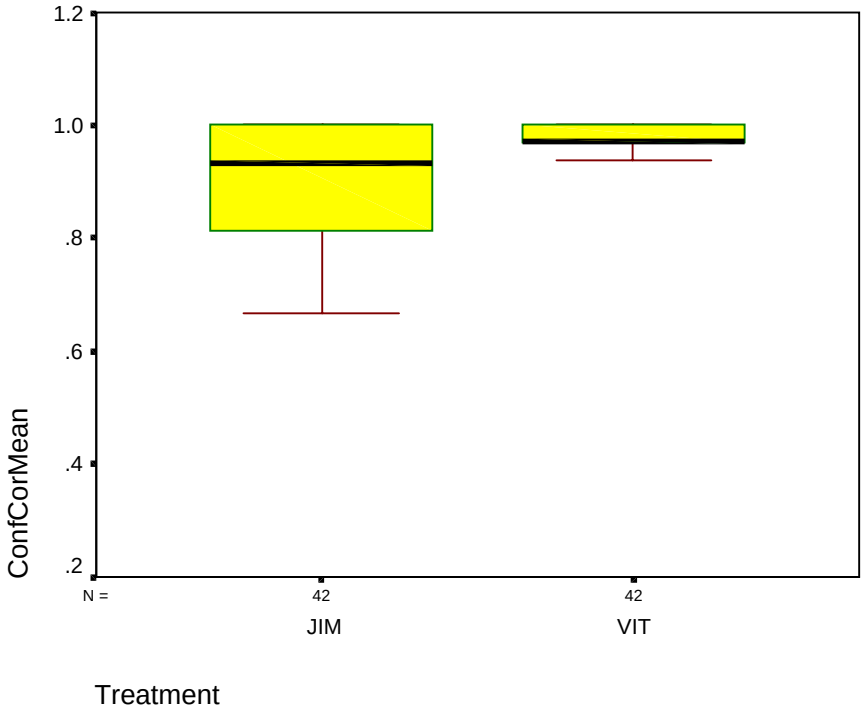


Figure 42 Confidence in Correctly Answered Tasks Box-plot

Examination of Table 14 and visual inspection of Figure 42 indicate that there is a difference between the two means and that the variance does not appear similar.

4.4.2 Hypothesis Test - Confidence

The observed significance level for Levene's test for equality of variance is less than 0.05 (L sig <<.001) so the *separate variance* t-test is used. The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean user confidence in correctly answered questions (t=-2.09, df=70, p<=.02). Therefore, the null hypothesis is rejected and the second research hypothesis is accepted:

H_{a2}: Subjects have greater confidence in their correctly answered tasks with the VITAMIN System than with the JIMI System.

Three additional null hypotheses are that, for each of the three task types, the mean confidence rating for the VITAMIN system is greater than or equal to the mean confidence rating for the JIMI system.

4.4.3 Type I Tasks - Confidence

The first three evaluated tasks were of task-type I. Descriptive statistics, box-plot, and hypothesis test are presented below for this task-type.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	.38	1	.90	.15
VITAMIN	42	.42	1	.92	.16

Table 15 Confidence in Correctly Answered Tasks – Type I

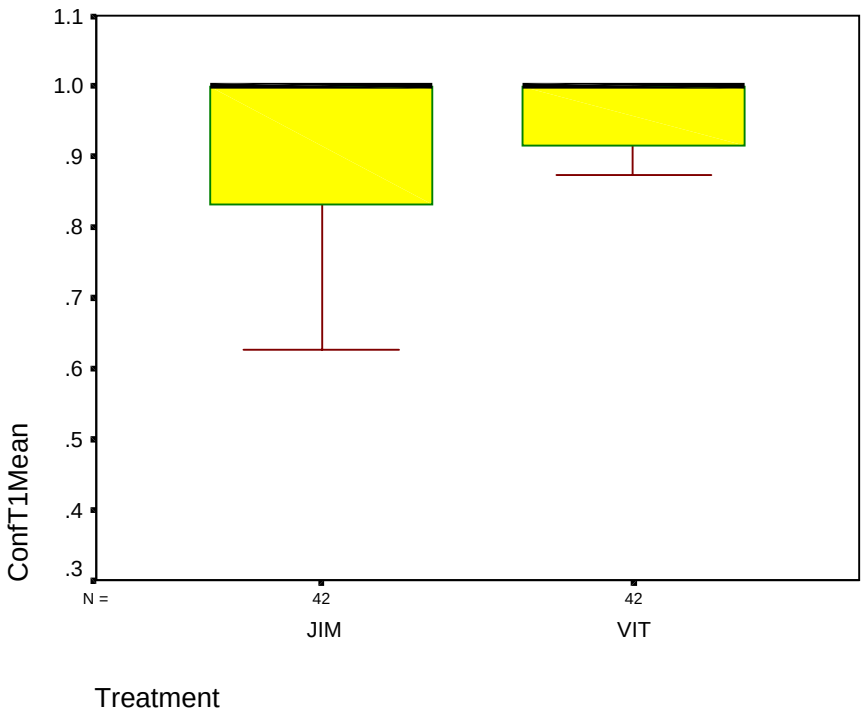


Figure 43 Confidence in Correctly Answered Tasks – Type I Box-plot

Examination of Table 15 and Figure 43 indicate almost no difference between the two means and that the variance appears similar; (L sig <<.001). The equality of variance

assumption is met, so the *pooled variance* t-test is used. The one-tailed independent samples *pooled variance* t-test found no statistically significant differences in the mean user confidence in correctly answered questions (t=-0.790, df=82, p<=.22). Therefore, the analysis fails to reject the null hypothesis for task-type I confidence.

4.4.4 Type II Tasks - Confidence

The second three evaluated tasks were of task-type II. Descriptive statistics, box-plot, and hypothesis test are presented below.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	0	1	.87	.23
VITAMIN	42	0	1	.93	.19

Table 16 Confidence in Correctly Answered Tasks – Type II

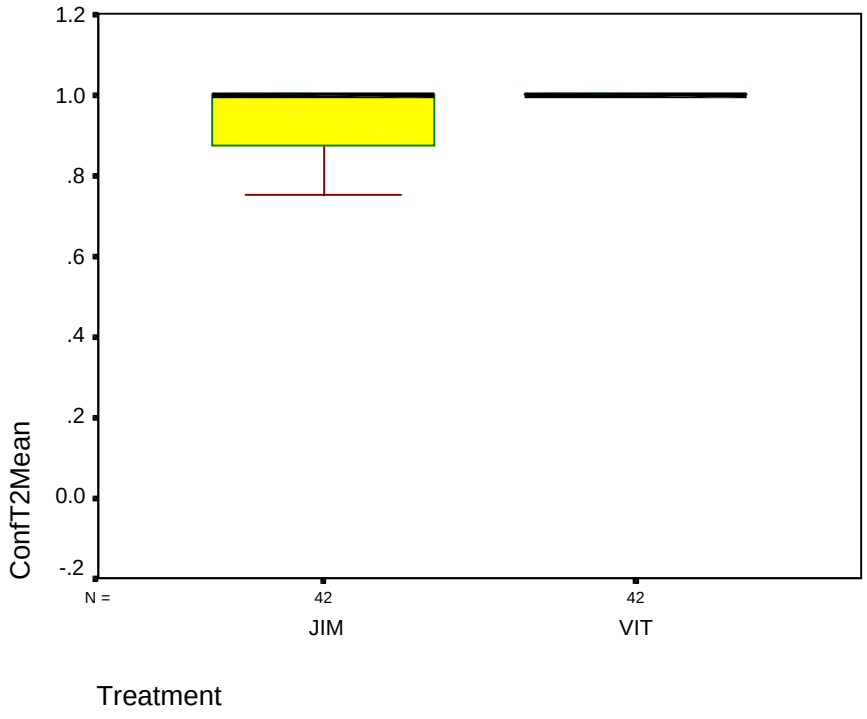


Figure 44 Confidence in Correctly Answered Tasks – Type II Box-plot

Examination of Table 16 and visual inspection of Figure 44 indicates almost no difference between the two means and that the variance appears similar; (L sig =.22). The *pooled variance* t-test found no statistically significant differences in the mean user confidence in correctly answered questions (t=-1.32, df=82, p<=.095). Therefore, the analysis fails to reject the null hypothesis for task-type II confidence.

4.4.5 Type III Tasks - Confidence

The third and last three evaluated tasks were of task-type III. Descriptive statistics, box-plot, and hypothesis test are presented below for this task-type.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	0	1	.85	.26
VITAMIN	42	.5	1	.94	.11

Table 17 Confidence in Correctly Answered Tasks – Type III

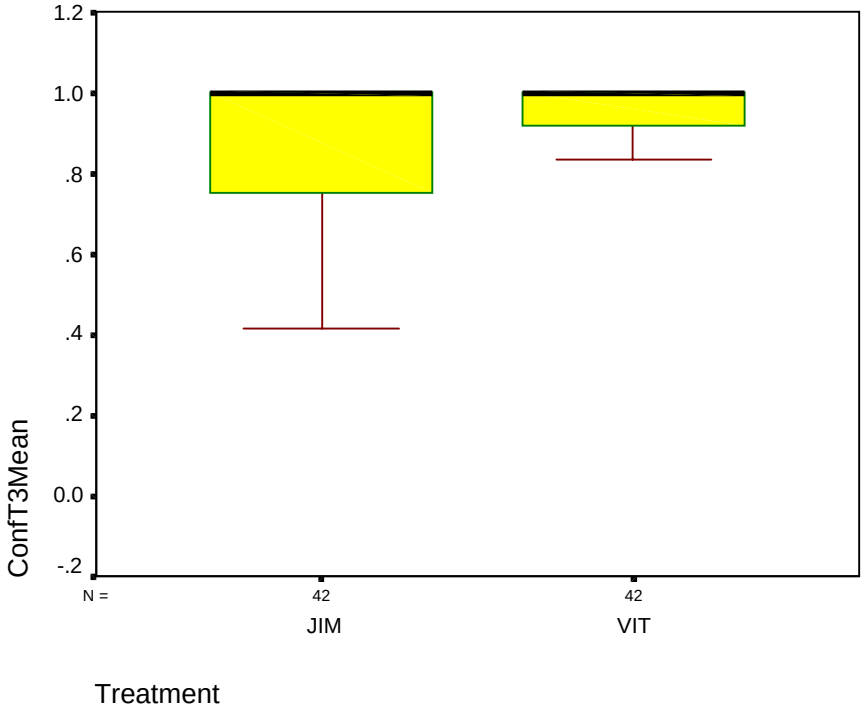


Figure 45 Confidence in Correctly Answered Tasks – Type III Box-plot

Examination of Table 17 and Figure 45 indicates almost no difference between the two means but the variance does not appear similar; (L sig <<.001). The equality of variance assumption is not met, so the *separate variance* t-test is used. The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean user confidence in correctly answered questions (t=-2.37, df=55, p<=.011). Therefore, the null hypothesis is rejected for task-type III confidence.

Overall, these three additional task hypotheses lend support to the advantage of the VITAMIN system over the JIMI system. There was almost no difference between the number of correct answers with the simplest Type I tasks (p<=.22). The Type II tasks showed differences that were not statistically significant (p<=.095 at alpha = .05), and the Type III tasks showed statistically significant differences (p<=.011).

4.4.6 Washout Assumption Test - Confidence

Application of the washout assumption test results in 42 cases, 23 when JIMI was the first treatment and 19 when VITAMIN was the first treatment as shown in Table 9.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	23	.33	1	.85	.18
VITAMIN	19	.61	1	.94	.11

Table 18 Confidence in Correctly Answered Tasks – Washout

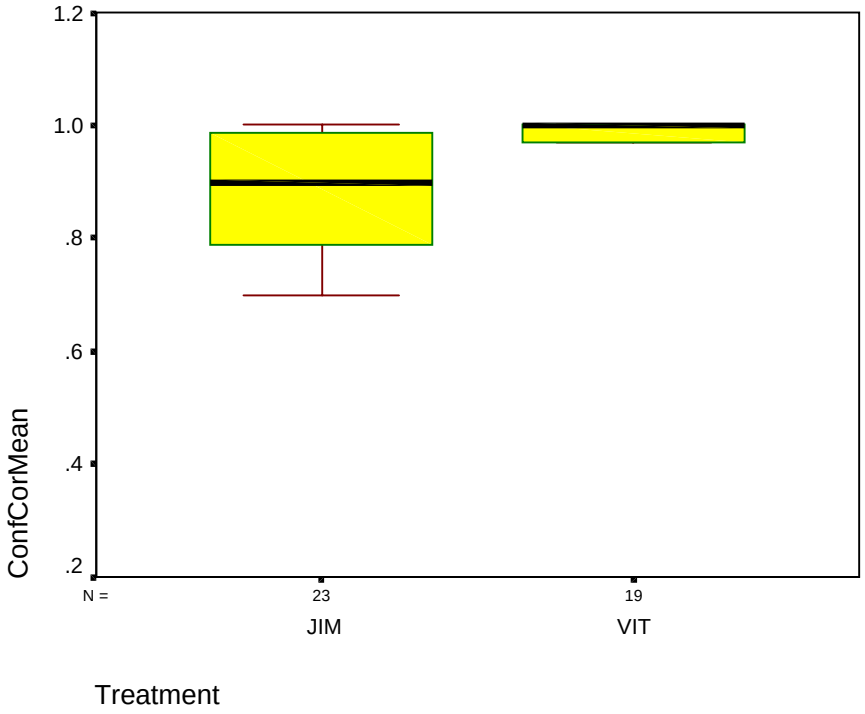


Figure 46 Confidence in Correctly Answered Tasks – Washout Box-plot

Examination of Table 18 and visual inspection of Figure 46 indicate that there is a difference between the two means and that the variance does not appear similar; however, L sig =.10). The equality of variance assumption is met, so the *pooled variance* t-test is used. The one-tailed independent samples *pooled variance* t-test found statistically significant differences in the mean user confidence in correctly answered questions (t=-2.1, df=40, p<=.023). Therefore, the null hypothesis is rejected and the second research hypothesis is still accepted. The washout assumption is valid for the second hypothesis.

4.5 Time To Answer Tasks Correctly

The *speed* factor corresponds to the average time the subject spent answering all questions correctly. This is calculated by subtracting the system query time, which

averages about 10 seconds per query from the time spent responding to the representative task. The time spent on each correctly answered question is summed and this total is divided by the number of correctly answered questions.

4.5.1 Descriptive Statistics - Speed

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	43	133	82	21
VITAMIN	42	20	64	34	8

Table 19 Time to Correctly Answer Tasks

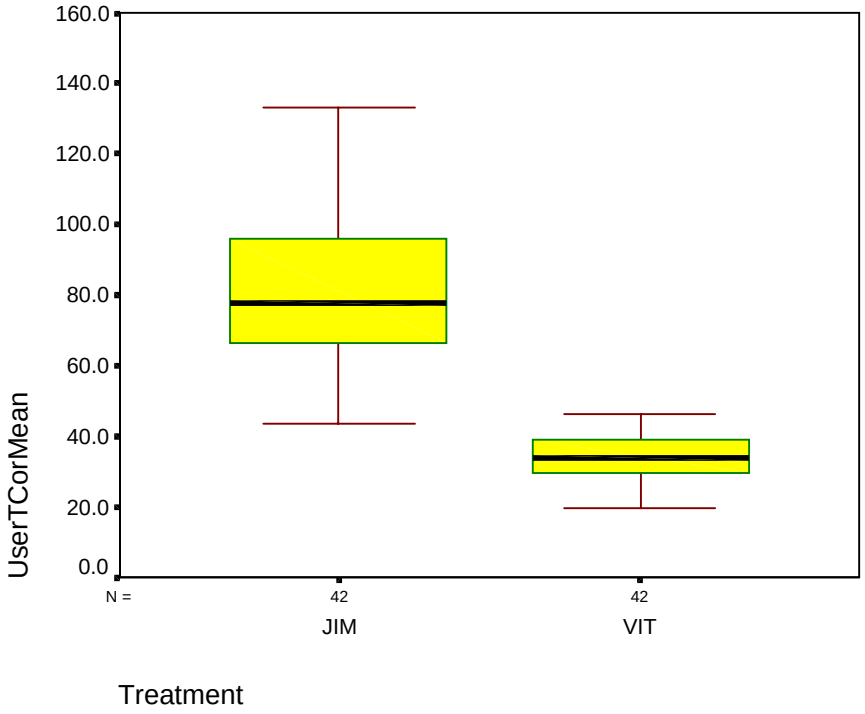


Figure 47 Time to Correctly Answer Tasks Box-plot

Examination of Table 19 and visual inspection of Figure 40 indicates that there is a large difference between the two means and that the variance does not appear similar. They barely overlap at the minimum and maximum.

4.5.2 Hypothesis Test - Speed

The observed significance level for Levene's test for equality of variance is less than 0.05 (L sig <<.001) so the *separate variance* t-test is used. The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean time to answer tasks correctly (t=13.8, df=53, p<<.001). Therefore, the null hypothesis is rejected and the third research hypothesis is accepted:

H_{a3}: Subjects answer tasks correctly in less time with the VITAMIN System than with the JIMI System.

Three additional null hypotheses are that, for each of the three task types, the mean time to answer tasks correctly for the VITAMIN system is less than or equal to the mean time to answer tasks correctly for the JIMI system.

4.5.3 Type I Tasks - Speed

The first three evaluated tasks were of task-type I. The descriptive statistics, box-plot, and hypothesis test are presented below.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	19	130	62	30
VITAMIN	42	9	69	28	14

Table 20 Time to Correctly Answer Tasks – Type I

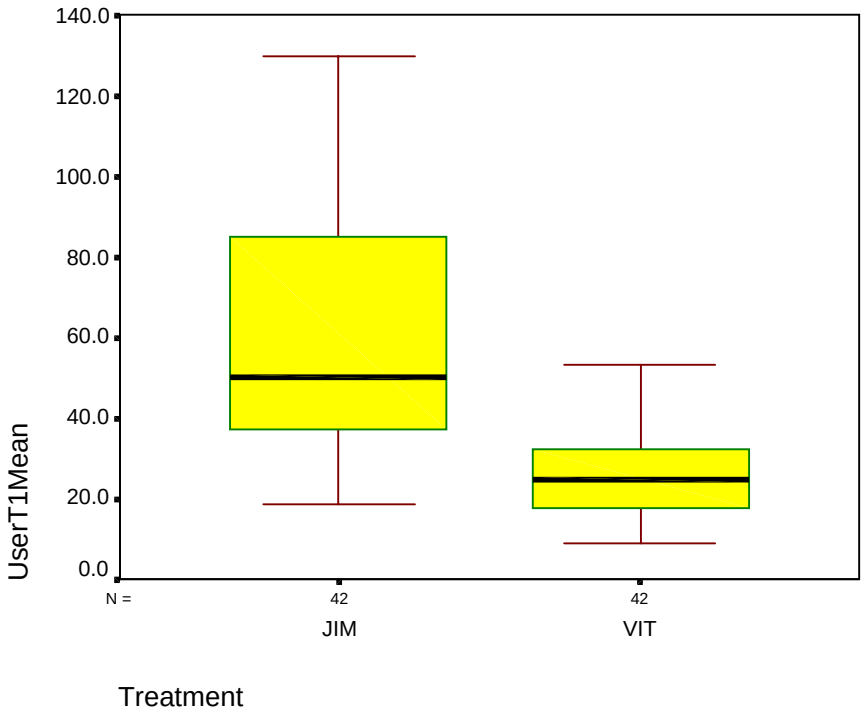


Figure 48 Time to Correctly Answer Tasks – Type I Box-plot

Examination of Table 20 and visual inspection of Figure 48 indicate differences between the two means and variances; (L sig <<.001). The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean time to answer tasks correctly (t=6.58, df=58, p<<.001). Therefore, the null hypothesis is rejected for task-type I speed.

4.5.4 Type II Tasks - Speed

The second three evaluated tasks were of task-type II. Descriptive statistics, box-plot, and hypothesis test are presented below.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	40	180	95	32
VITAMIN	42	16	53	29	10

Table 21 Time to Correctly Answer Tasks – Type II

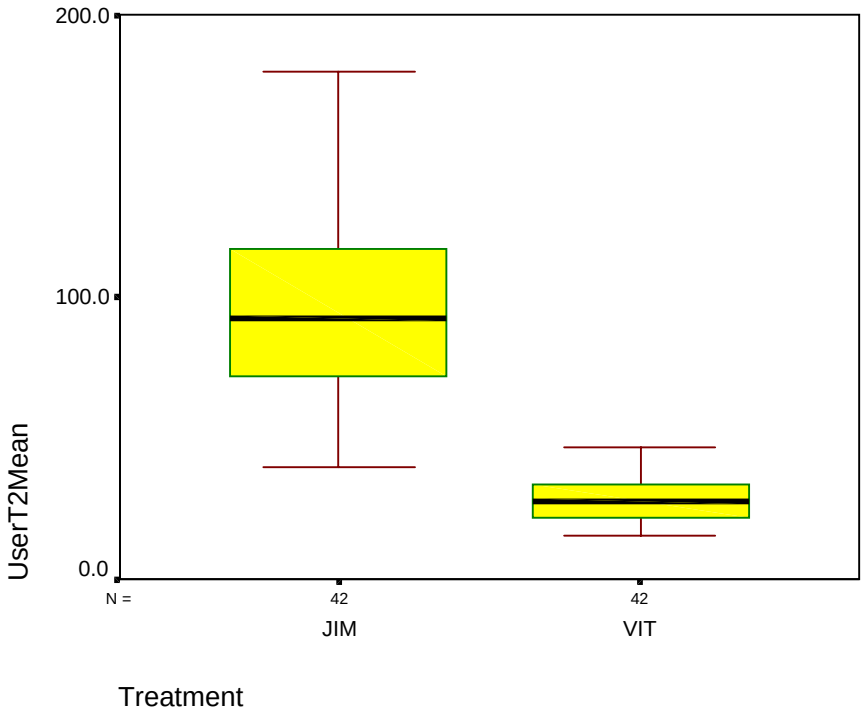


Figure 49 Time to Correctly Answer Tasks – Type II Box-plot

Examination of Table 21 and visual inspection of Figure 49 indicate differences between the two means and variances; (L sig <<.001). The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean time to answer tasks correctly (t=12.7, df=49, p<<.001). Therefore, the null hypothesis is rejected for task-type II speed.

4.5.5 Type III Tasks - Speed

The third and last three evaluated tasks were of task-type III. Descriptive statistics, box-plot, and hypothesis test are presented below for this task-type.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	30	232	93	38
VITAMIN	42	24	85	46	13

Table 22 Time to Correctly Answer Tasks – Type III

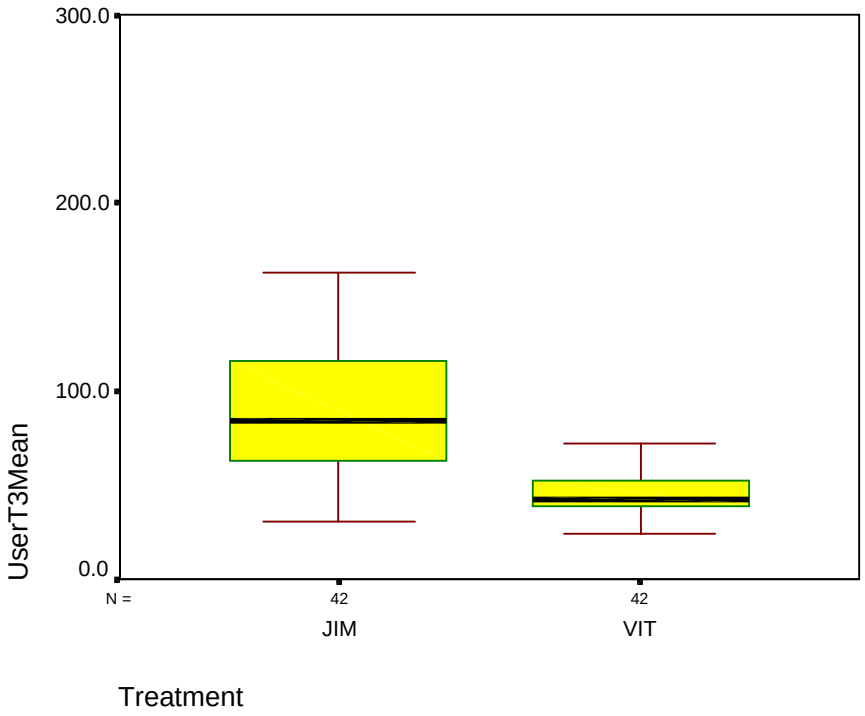


Figure 50 Time to Correctly Answer Tasks – Type III Box-plot

Examination of Table 22 and visual inspection of Figure 50 indicate differences between the two means and variances; (L sig <<.001). The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean time to answer tasks correctly (t=7.56, df=50, p<<.001). Therefore, the null hypothesis is rejected for task-type III speed.

Overall, these three additional task hypotheses lend support to the advantage of the VITAMIN system over the JIMI system in speed. All three task-types showed statistically significant differences with p<<.001. These results indicate that, as expected, VITAMIN was more helpful for task-type II than for task-type I. However, the advantage was mitigated somewhat with task-type III. Possible reasons for this disparity are addressed in Chapter 5; in short, several subjects discovered that a single, precise query issued from the JIMI system could answer the representative task.

4.5.6 Washout Assumption Test - Speed

Discarding the second treatment results in 42 cases, 23 when JIMI was the first treatment and 19 when VITAMIN was the first treatment, as shown in Table 12.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	23	51	133	86	23
VITAMIN	19	30	64	39	8

Table 23 Time to Correctly Answer Tasks – Washout

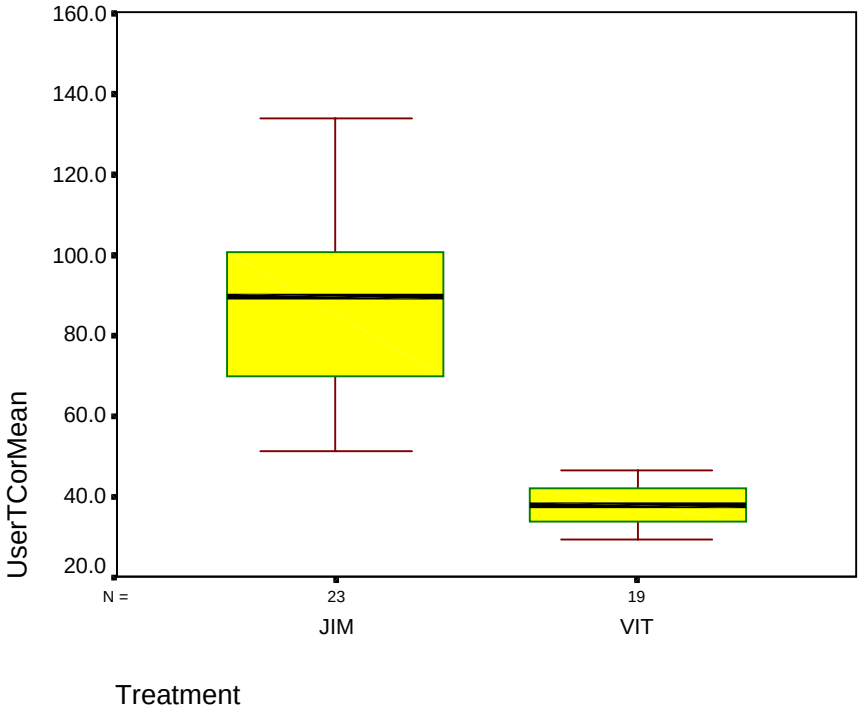


Figure 51 Time to Correctly Answer Tasks – Washout Box-plot

Examination of Table 23 and visual inspection of Figure 51 indicate differences between the two means and variances; (L sig <<.001). The one-tailed independent samples *separate variance* t-test found statistically significant differences in the mean time to answer tasks correctly (t=9.35, df=28, p<<.001). Therefore, the null hypothesis is rejected and the third research hypothesis is still accepted. The washout assumption is valid for the third hypothesis, speed, or time to correctly answer tasks.

4.6 User Satisfaction With The Interface

The *satisfaction* factor corresponds to the subject's overall satisfaction rating for the system under consideration. This total satisfaction score is the sum of the scores for each attribute as shown in Table 24. Additionally, the score and analysis is provided for each of the eight attributes and the survey reliability analysis is reported.

Multiple Choice Selection	Analysis Value	Sum Analysis Value
A. Strongly Disagree	0.00	0.0
B. Disagree	0.25	2.0
C. Neither Agree no Disagree	0.50	4.0
D. Agree	0.75	6.0
E. Strongly Agree	1.00	8.0

Table 24 Survey Analysis Values

4.6.1 Descriptive Statistics - Satisfaction

For overall satisfaction, 0.0 is strong dissatisfaction, 4.0 is neutral, and 8.0 is strong satisfaction. The unusual Likert scale is due to the summing of the eight attributes as illustrated in Table 24. These results are shown in Table 25.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	42	0	6	3.4	1.7
VITAMIN	42	1	8	5.7	1.5

Table 25 User Satisfaction with the IMIS

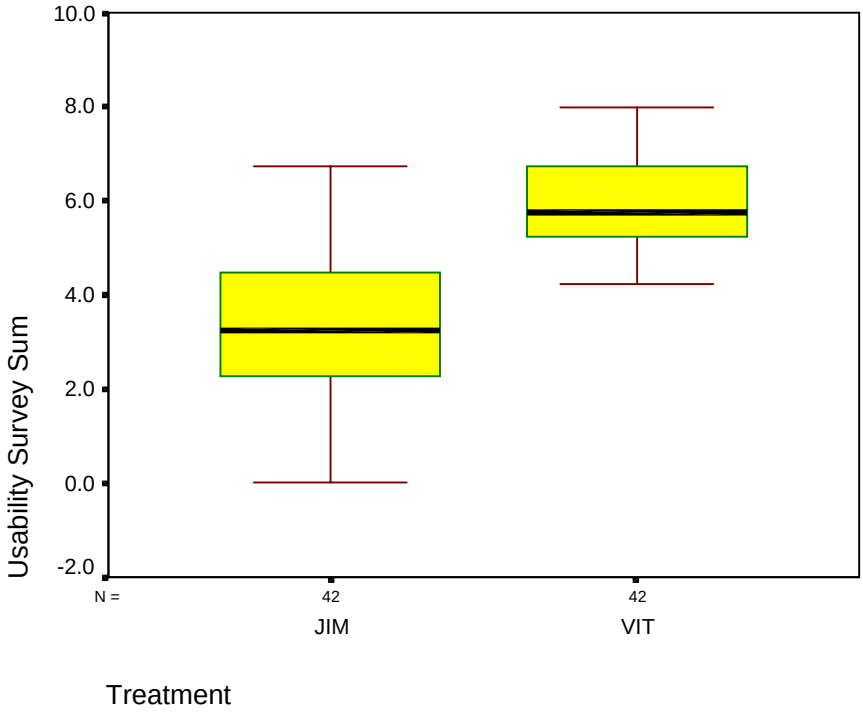


Figure 52 User Satisfaction with the IMIS Box-plot

Examination of Table 25 and visual inspection of Figure 52 indicate that there is a difference between the two means and that the variance appears similar.

4.6.2 Hypothesis Test - Satisfaction

The observed significance level for Levene's test for equality of variance is greater than 0.05; (L sig =.098). The equality of variance assumption is met, so the *pooled variance* t-test is used. The one-tailed independent samples *pooled variance* t-test found statistically significant differences in the mean number of correctly answered questions (t=-6.57, df=82, p<<.001). Therefore, the null hypothesis is rejected and the fourth research hypothesis, overall satisfaction, is accepted:

H_{a4}: Subjects are more satisfied with the VITAMIN System than with the JIMI System.

4.6.3 Washout Assumption Test - Satisfaction

Application of the washout assumption test results in 42 cases, 23 when JIMI was the first treatment and 19 when VITAMIN was the first treatment as shown in Table 9.

	N	Minimum	Maximum	Mean	Std. Deviation
JIMI	23	1	6	3.3	1.5
VITAMIN	19	1	8	5.8	1.5

Table 26 User Satisfaction with the IMIS – Washout

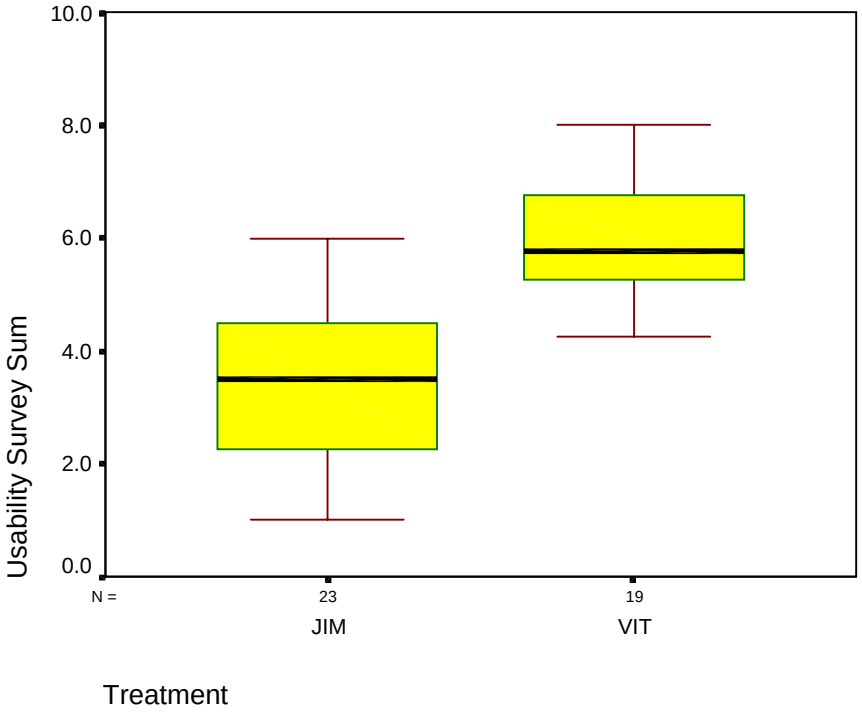


Figure 53 User Satisfaction with the IMIS – Washout Box-plot

Examination of Table 26 and visual inspection of Figure 53 indicate that there is a difference between the two means and that the variance appears similar; (L sig =.43) so the *pooled variance* t-test is used. The one-tailed independent samples *pooled variance* t-test found statistically significant differences in the overall satisfaction rating (t=-5.23, df=40, p<<.001). Therefore, the null hypothesis is rejected and the fourth research hypothesis is still accepted. The washout assumption is valid for the fourth hypothesis.

4.6.4 Individual Satisfaction Measures

This section presents the analysis for each factor in the satisfaction instrument individually. The descriptive statistics and hypothesis tests are presented in Tables 27 and 28, then the box-plots and analysis are reported in Figures 54 through 61.

Factor	JIMI Mean	JIMI StDev	VITA Mean	VITA StDev
The system was easy to use.	0.35	0.30	0.75	0.30
The system was helpful.	0.52	0.28	0.69	0.25
The system allowed me to perform faster.	0.32	0.34	0.75	0.29
The system provided high information quality.	0.57	0.30	0.74	0.23
The system provided high interface quality.	0.38	0.22	0.74	0.21
The system allowed me to learn about the data.	0.48	0.26	0.71	0.22
The system was enjoyable to use.	0.26	0.27	0.66	0.27
The system was useful.	0.54	0.29	0.71	0.23

Table 27 User Satisfaction Individual Measures - Summary

Factor	L sig	t-value	df	p-value t_α	significant
The system was easy to use.	.14	-6.04	82	<<.001	yes
The system was helpful.	.064	-2.87	82	.002	yes
The system allowed me to perform faster.	.21	-6.23	82	<<.001	yes
The system provided high information quality.	**.010**	-2.82	77	.003	yes
The system provided high interface quality.	.21	-7.76	82	<<.001	yes
The system allowed me to learn about the data.	**.031**	-4.38	79	<<.001	yes
The system was enjoyable to use.	.78	-6.87	82	<<.001	yes
The system was useful.	**.013**	-3.07	78	.001	yes

Table 28 User Satisfaction Individual Measures - Hypotheses

The additional hypotheses analyzed in this section are that subjects are more satisfied with the VITAMIN System than with the JIMI System for each of the satisfaction factors.

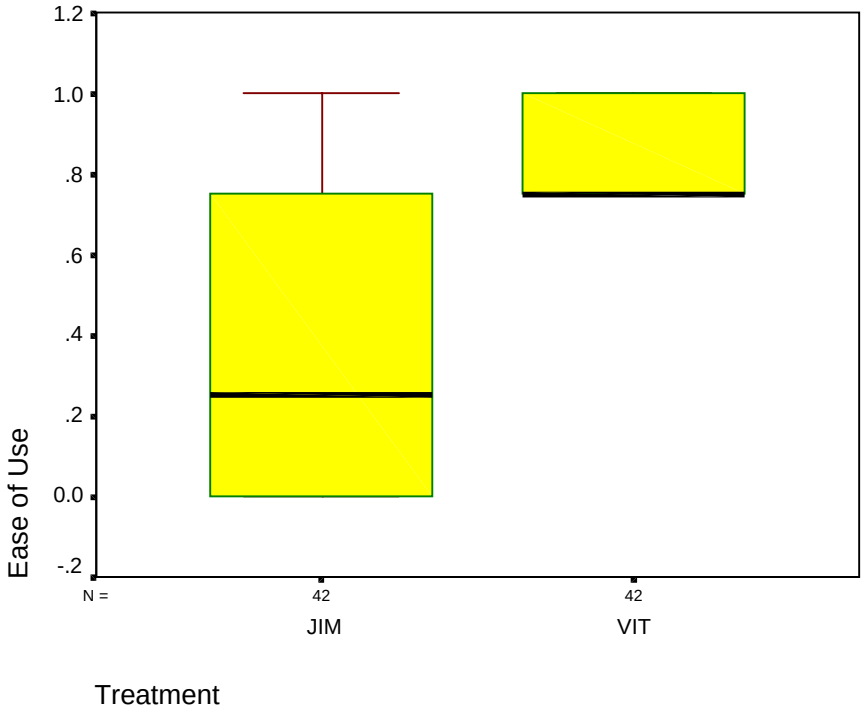


Figure 54 The System Was Easy To Use. Box-plot

Examination of Table 27 and visual inspection of Figure 54 indicate that there is a difference between the two means and that the variance does not appear similar. Table 28 contains the results of the Levene and hypothesis tests. The observed significance level for Levene's test for equality of variance is greater than 0.05. The one-tailed independent samples *pooled variance* t-test found statistically significant differences for the ease of use rating as shown in Table 28.

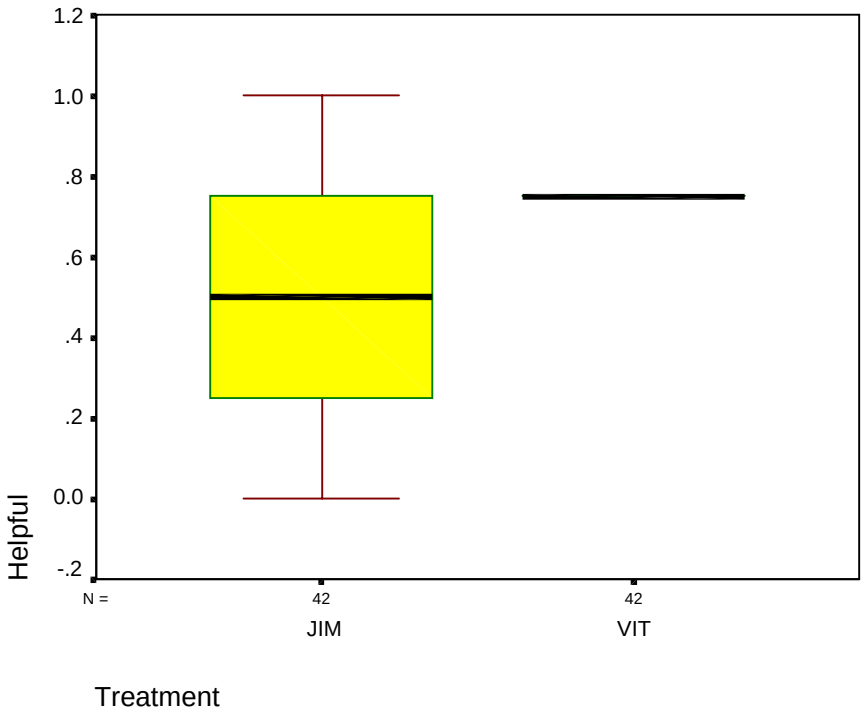


Figure 55 The System Was Helpful. Box-plot

Examination of Table 27 and visual inspection of Figure 55 indicate that there is a difference between the two means and that the variance does not appear similar. Table 28 contains the results of the Levene and hypothesis tests. The observed significance level for Levene's test for equality of variance is greater than 0.05. The one-tailed independent samples *pooled variance* t-test found statistically significant differences for the helpfulness rating as shown in Table 28.

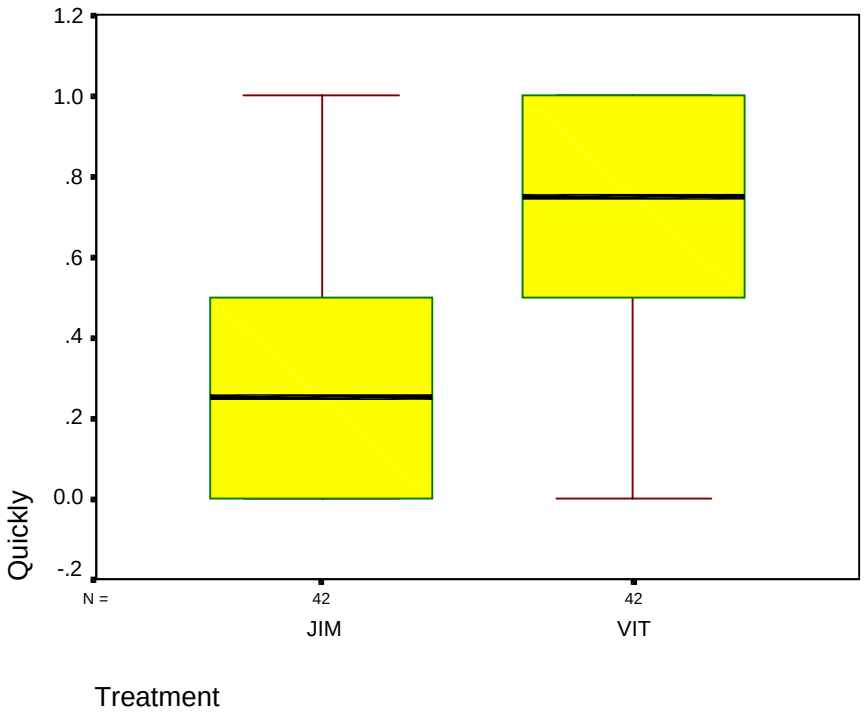


Figure 56 The System Allowed Me To Perform Faster. Box-plot

Examination of Table 27 and visual inspection of Figure 56 indicate that there is a difference between the two means and that the variance does not appear similar. Table 28 contains the results of the Levene and hypothesis tests. The observed significance level for Levene's test for equality of variance is greater than 0.05. The one-tailed independent samples *pooled variance* t-test found statistically significant differences for the speed rating as shown in Table 28.

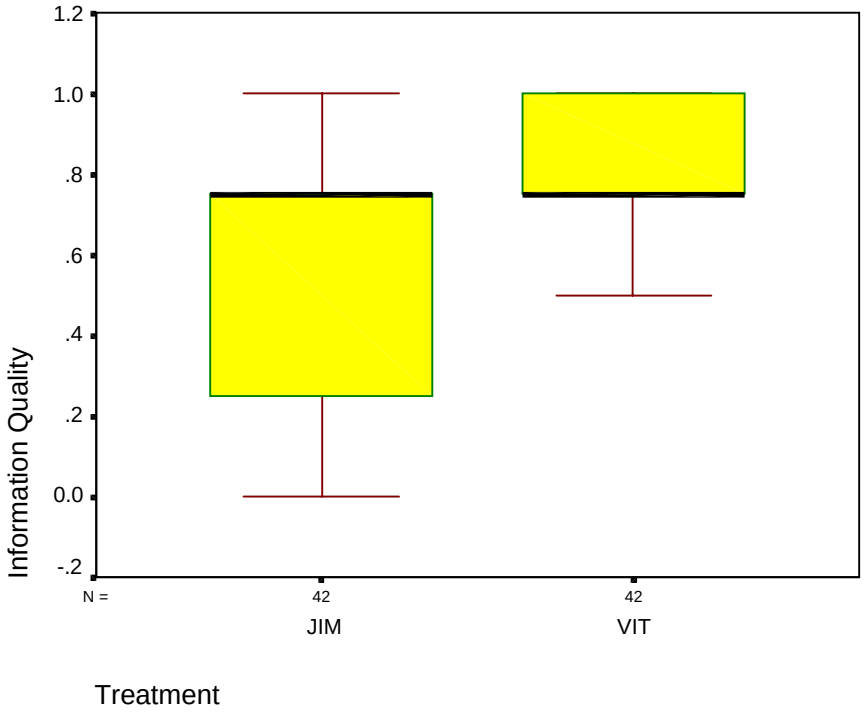


Figure 57 The System Provided High Information Quality. Box-plot

Examination of Table 27 and visual inspection of Figure 57 indicate that there is a difference between the two means and that the variance does not appear similar. Table 28 contains the results of the Levene and hypothesis tests. The observed significance level for Levene's test for equality of variance is less than 0.05. The one-tailed independent samples *separate variance* t-test found statistically significant differences for the information quality rating as shown in Table 28.

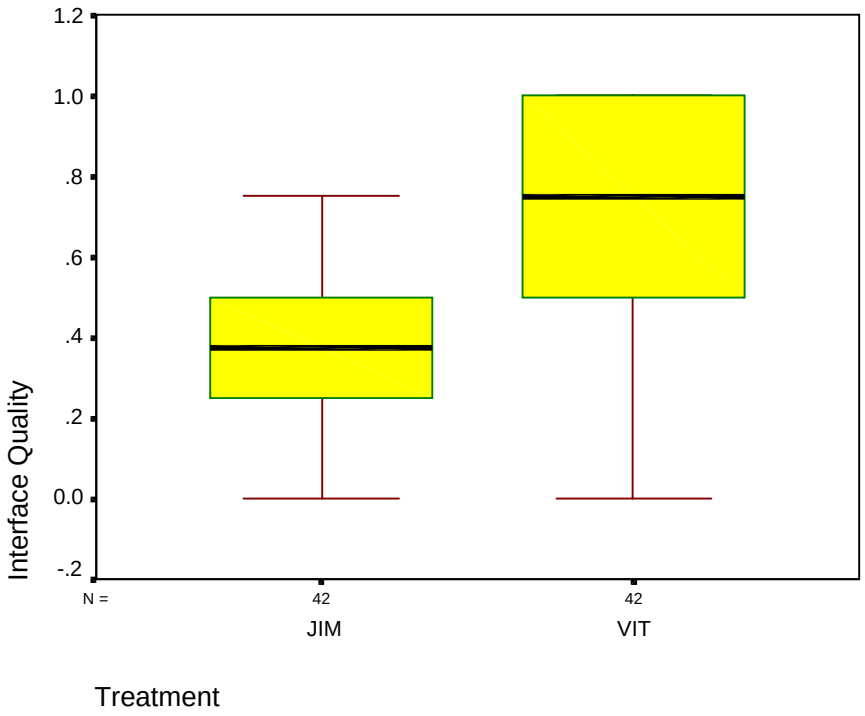


Figure 58 The System Provided High Interface Quality. Box-plot

Examination of Table 27 and visual inspection of Figure 58 indicate that there is a difference between the two means and that the variance does not appear similar. Table 28 contains the results of the Levene and hypothesis tests. The observed significance level for Levene's test for equality of variance is greater than 0.05. The one-tailed independent samples *pooled variance* t-test found statistically significant differences for the interface quality rating as shown in Table 28.

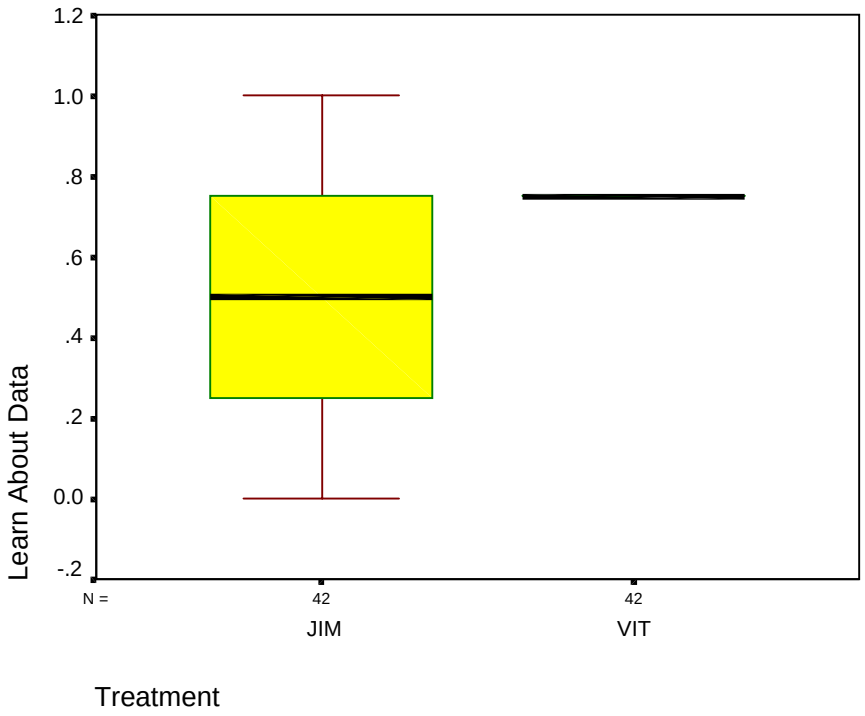


Figure 59 The System Allowed Me To Learn About The Data. Box-plot

Examination of Table 27 and visual inspection of Figure 59 indicate that there is a difference between the two means and that the variance does not appear similar. Table 28 contains the results of the Levene and hypothesis tests. The observed significance level for Levene's test for equality of variance is less than 0.05. The one-tailed independent samples *separate variance* t-test found statistically significant differences for the 'learn about the data' rating as shown in Table 28.

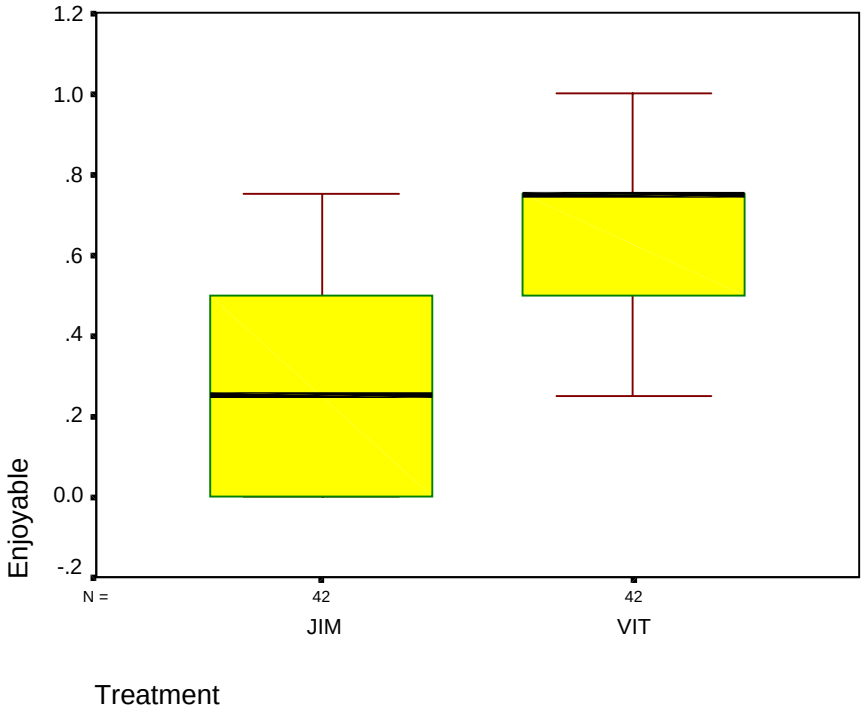


Figure 60 The System Was Enjoyable To Use. Box-plot

Examination of Table 27 and visual inspection of Figure 60 indicate that there is a difference between the two means and that the variance does not appear similar. Table 28 contains the results of the Levene and hypothesis tests. The observed significance level for Levene's test for equality of variance is greater than 0.05. The one-tailed independent samples *pooled variance* t-test found statistically significant differences for the enjoyability rating as shown in Table 28.

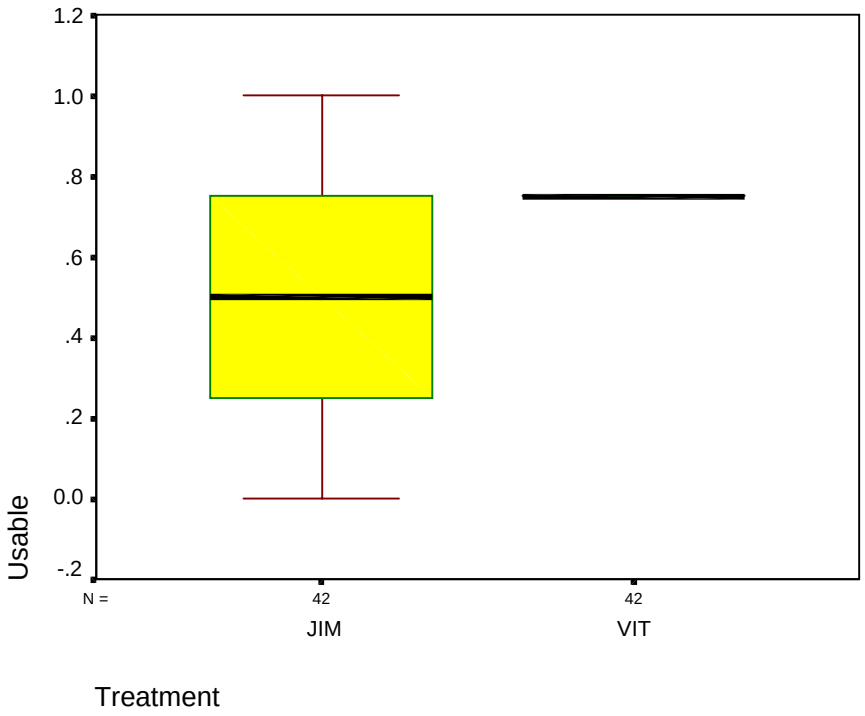


Figure 61 The System Was Useful. Box-plot

Examination of Table 27 and visual inspection of Figure 61 indicate that there is a difference between the two means and that the variance does not appear similar. Table 28 contains the results of the Levene and hypothesis tests. The observed significance level for Levene's test for equality of variance is less than 0.05. The one-tailed independent samples *separate variance* t-test found statistically significant differences for the usability rating as shown in Table 28.

4.6.5 Reliability Analysis - Satisfaction

Reliability analysis includes an item analysis to determine whether the attitude scale is measuring one or more dimensions and an analysis of the discriminatory power of each statement. The results of the last section show a reasonable correlation among the survey questions. This item analysis demonstrates that the attitude scale is measuring one dimension, satisfaction. The second analysis measures the discriminatory power as described by Clark-Carter [CLA97].

For the single dimension of satisfaction measured by this instrument, the following procedure was used to analyze the discriminatory power:

1. The sum of the responses for each statement was calculated to find total score of each subject. These results ranged from zero to eight as reported in the previous section.

2. The subjects who gave the top 25% and the bottom 25% of the total scores were identified. Twenty-one subjects were in each group from the total of 84. The bottom 25% ranged in ratings from 0.0 to 3.0 and the top 25% ranged from 6.0 to 8.0.

3. For each statement, the responses in these two groups were summed.

4. The statement sums were compared to determine whether these two groups differed in their responses. These results are reported in the table below.

Statements that fail to distinguish between subjects who give high scores and those who give low scores have poor discriminative power. Such statements are eliminated from the instrument by this analysis. In the case of this instrument, all statements strongly distinguished between these two groups as shown in Table 29.

Factor	Bottom 25%	Top 25%
The system was easy to use.	3.00	19.75
The system was helpful.	5.00	17.25
The system allowed me to perform faster.	3.75	20.00
The system provided high information quality.	8.00	18.50
The system provided high interface quality.	5.75	17.50
The system allowed me to learn about the data.	5.25	17.50
The system was enjoyable to use.	0.75	15.75
The system was useful.	5.75	17.50

Table 29 Discriminatory Analysis

4.7 Qualitative Results

Examining the context of quantitative data is an important element of usability inquiry. Statistical portrayals, especially for remote evaluations, must be interpreted and given meaning. To assist in this analysis, open-ended survey responses were requested of all subjects. The final three survey questions requested observations from the subjects concerning VITAMIN, JIMI, and overall comments. These open-ended questions capture the subjects' responses in their own terms. The purpose of these free-text questions was to provide qualitative richness and depth to the predominantly quantitative data. These responses allow a thorough examination of the context to enable the discovery of additional factors that were not requested by the experimental design. The open-ended questions also facilitate the identification of themes that may substantiate the quantitative analysis. All forty-one subjects provided commentary; these raw unedited responses are provided in Appendix K.

This section analyzes selected extracts from this free-text survey data. Overall, most subjects indicated a preference for the VITAMIN system. Thirty subjects preferred VITAMIN, five preferred JIMI, and six had no clear preference. Initial analysis revealed that the qualitative themes emerging from the subjects' commentary corresponded to factors of the coherence metric. The next four sections present illustrative extracts from these comments. These sections are organized by coherence factors of the indirect manipulation interface system (IMIS), correctness, confidence, speed, and satisfaction (C^2S^2). Following these four sections, a qualitative analysis section summarizes the themes expressed by the subjects. Each new paragraph represents a different subject in the comments presented below.

4.7.1 Correctness

Somehow, it seemed like a computer game where you are searching for something and you have to go into doors to find. I found VITAMIN to be very enjoyable to work with and. **never got frustrated or felt my answers were incorrect**

For soldiers on the battle field, it is **more convenient** for them to push a button and get the information they need rather

than sitting there trying to type in information which leaves room for errors.

The main thing that I liked about JIMI was that it was **very simple to compare** the data of the different ships, companies, and other components once they were drawn up.

4.7.2 Confidence

I was **more sure** that the information that I had retrieved was right. [VITAMIN]

After the first few questions, **I felt very confident** in how to use the [JIMI] system and I enjoyed it greatly.

4.7.3 Speed

I moved through the questions at a **much faster rate** with the VITAMIN system.

The [VITAMIN] system allowed me to do the **work faster**.

Excellent interface. Plus and minus buttons made it **much faster and easier to use**. [VITAMIN]

- VITAMIN's interface proved only to **slow me down** as the problems became more complicated.

- **[VITAMIN] was difficult, painstaking, and time-consuming** to compare the different companies and ships to one another. [It was] extremely simple to use and find answers. However, once those answers were drawn up it was a little bit more of a difficult process, causing me to constantly scroll up and scroll down.

JIMI was faster and gave all of the data up front. I think the JIMI system would be the better system after knowing what the data meant.

I feel as though I could get more information off of the **JIMI** system. I first needed to find out which place I had to go to retrieve the correct information, and then I needed to be sure to put the words in the correct place. Once I was able to do that, **the system worked faster**.

4.7.4 Satisfaction

Much easier to use, organize, and tell what is going on enough to answer the questions without having a real good computer understanding and background.

I liked the fact that all of the information was displayed on the screen. Clicking on the plus icons made it **easy to access more data**.

The [VITAMIN] system **allowed me to comprehend the material easier** because it was in a **logical chart format that I am familiar with**.

- **The JIMI was very complicated to use**. For someone like myself to use the JIMI it is very frustrating. It is also time consuming because it required you to type in the information you were looking for.

- The JIMI system was **very hard to use**. The information was hard to find and errors were made more frequently. The information was easily read, but it was hard to find. The information **was also slow** to come up.

- [JIMI] was **more difficult to use and more tedious**. The problem is that it did take me a little longer to input the words (especially since it was easier to make a mistake). It was not as user friendly.

4.7.5 Overall Results

[VITAMIN] was a very nice system in that it provided a visual classification that was **easy to comprehend**.

VITAMIN would be **much better for new workers** because it is very easy to learn and understand.

The [VITAMIN] system **allowed me to comprehend the material easier because it was in a logical chart format that I am familiar with**.

I think **the JIMI system would be the better system after knowing what the data meant**.

I feel as though **I could get more information off of the JIMI system**. I first needed to find out which place I had to go to retrieve the correct information, and then I needed to be sure to put the words in the correct place.

JIMI requires time to think about what is what are variables and their meanings.

- The JIMI system was very confusing. [It was] difficult to find the data needed because it was **not laid out clearly**.

- I disliked the JIMI format. I was **very cumbersome** and annoying to have to type in different answers to search for every time. It was also difficult to understand what some of the symbols stood for.

- I found [JIMI] confusing to use. It was **more complicated** than the VITAMIN program, and took me longer to use. I didn't like the fact that not all of the information was visible at one time. This made it difficult to compare data quickly.

- The JIMI was **very complicated** to use. For someone like myself to use the JIMI it is very frustrating. It is also time consuming because it required you to type in the information you were looking for.

4.7.6 Qualitative Analysis

Several themes emerge from the analysis of this qualitative data. First, nearly all subjects found the VITAMIN IMIS easier to use and easier to learn. VITAMIN was designed with visual cues that allow the data to be presented more coherently. Many of the subject comments indicate an intuitive appreciation for the enhancements of the VITAMIN system. These visual cues include improved data selection instead of data entry, improved visual clarity with familiar color effects, and improved representation with an aggregating tree structure. In their own terms, the subjects identified these enhancements with descriptions such as the "red and green plusses" or "colored buttons" or "doors" that allow point and select aggregation and disaggregation of the agent-mediated data in an "outline".

Most subjects reported great satisfaction with the VITAMIN system and found it faster (*speed*), "easy to comprehend" (*satisfaction*), and able to generate results in which they had great *confidence*. The subjects that expressed preferences for the JIMI system typically qualified the preference by addressing the learnability of VITAMIN. An example of such a response from a subject that indicated some preference for JIMI is this reply:

> Both of these systems are accurate and quite effective. I did not find either of them to be difficult to use. With the proper tutorial and training, an individual could use either program. For the sake of simplicity and lack of complication VITAMIN would be more effective, however with a more experienced individual looking for expediency of data JIMI would seem to me to be the ideal program.

The subjects that preferred the JIMI interface generally considered that it was faster and allowed them to "go directly to the data." They considered the JIMI system's "higher learning curve" to be time "well invested." Several of the subjects also preferred the simple manner in which the JIMI system displayed the results. Interestingly, according to Borenstein [BOR94], these findings are similar to comments typically attributed to expert users. Expert users often express concern when "enhancements" appear to impede access to information. The JIMI system's text-entry design allowed knowledgeable users

to specify actions in several dimensions with a single query. Experts typically prefer interface systems that enable "shortcuts." For instance, task-type III tasks could be answered with a few carefully crafted queries instead of many single dimension queries required by the VITAMIN system. There was little demographic difference among the pool of subjects so these anomalies represent individual differences. The qualitative responses lead to verification of this 'carefully crafted query' supposition for several subjects from the by-question quantitative data in Table 37 of Appendix J and in Table 51 of Appendix L. This effect is especially noticeable for question 10 where several subjects issued far fewer queries and answered the task in far less time with JIMI than with VITAMIN.

Subjects reported many more negative valuation comments for the JIMI system than for the VITAMIN system. Most subjects reported the JIMI was "confusing," and "hard to use." They were "unsure about" the answers and did not enjoy typing the queries. Negative comments about VITAMIN generally included the perception that it was slower and that the aggregating display was difficult to navigate. Less than one fifth of the qualitative comments indicated a preference for the JIMI system.

Several recurring themes emerge from the qualitative data identified above. These themes include confidence, speed, and satisfaction. This qualitative data and these themes support the quantitative data analysis and generate strong resultant support for the concept of coherence introduced in this research.

4.8 Analysis Summary

Statistically significant results were obtained for all four indicators of coherence represented by the four research hypotheses as shown in Table 30. The enhanced coherence of the VITAMIN system afforded an improvement over JIMI for: (1) the number of correctly answered tasks; (2) the user confidence in correctly answered tasks; (3) the time to answer tasks correctly (speed); and (4) the user satisfaction with the interface overall and for each satisfaction factor.

Hypothesis Factor	JIMI Mean	JIMI StDev	VITA Mean	VITA StDev	p-value (t_α)
H_{01} - Correctness	7.8	1.4	8.6	0.8	<=.002
H_{02} - Confidence	0.87	0.18	0.94	0.12	<=.020
H_{03} - Speed	82	21	34	8	<<.001
H_{04} - Satisfaction	3.4	1.7	5.7	1.5	<<.001

Table 30 Overall Results

Additional support was given by the three additional task-type hypotheses in that VITAMIN provided increased support for the more difficult tasks. In addition, the washout assumption held for all hypotheses. Additional analysis is reported in Appendix L. This analysis includes tests of the washout assumption by task type, tests of the hypothesis by each individual question, and tests of the washout assumption by questions. The overall summary results of these additional analyses are presented in Table 31 and Table 32.

Hypothesis Factor	Task-Type	CBD (t_α)	CBD Reject H_0	Washout (t_α)	Washout Reject H_0	Delta
Correctness	I	<=0.216	no	<=0.094	no	no
Confidence	I	<=0.264	no	<=0.122	no	no
Speed	I	<<0.001	yes	<<0.001	yes	no
Correctness	II	<=0.059	no	<=0.029	yes	yes
Confidence	II	<=0.095	no	<=0.103	no	no
Speed	II	<<0.001	yes	<<0.001	yes	no
Correctness	III	<<0.001	yes	<=0.008	yes	no
Confidence	III	<=0.011	yes	<=0.011	yes	no
Speed	III	<<0.001	yes	<<0.001	yes	no

Table 31 Results by Task-Type

The results of the counterbalanced design (CBD) indicate increasing support for the coherent interface with increasing task-type complexity. With task-type I, only one research hypothesis was accepted. Task-type II had two accepted research hypotheses. All three null hypotheses were rejected and therefore, all three research hypotheses accepted for the most complex type III tasks. The washout results are identical except that the number of tasks answered correctly was also significant for task-type II; that is, the null hypothesis was also rejected and this was different from the CBD as indicated in the "Delta" column in Table 31. Therefore, the washout assumption test resulted in a stronger effect than the CBD. Finally, Table 32 presents the overall results of the by-question analysis.

Hypothesis Factor	Question	CBD (t_α)	CBD Reject H_0	Washout (t_α)	Washout Reject H_0	Delta
Correctness	4	<=0.375	no	<=0.172	no	no
Confidence	4	<=0.089	no	<=0.026	yes	yes
Speed	4	<<0.001	yes	<=0.001	yes	no
Correctness	5	constant	no	constant	no	no
Confidence	5	<=0.439	no	<=0.454	no	no
Speed	5	<=0.001	yes	<=0.018	yes	no
Correctness	6	<=0.148	no	<=0.109	no	no
Confidence	6	<=0.039	yes	<=0.007	yes	no
Speed	6	<<0.001	yes	<<0.001	yes	no
Correctness	7	<=0.022	yes	<=0.042	yes	no
Confidence	7	<=0.014	yes	<=0.007	yes	no
Speed	7	<<0.001	yes	<<0.001	yes	no
Correctness	8	<=0.156	no	<=0.081	no	no
Confidence	8	<=0.047	yes	<=0.059	no	yes
Speed	8	<<0.001	yes	<<0.001	yes	no
Correctness	9	<=0.364	no	<=0.338	no	no
Confidence	9	<=0.028	yes	<=0.219	no	yes
Speed	9	<<0.001	yes	<<0.001	yes	no
Correctness	10	<=0.325	no	<=0.226	no	no
Confidence	10	<=0.124	no	<=0.051	no	no
Speed	10	<=0.477	no	<=0.166	no	no
Correctness	11	<<0.001	yes	<=0.011	yes	no
Confidence	11	<<0.001	yes	<=0.002	yes	no
Speed	11	<<0.001	yes	<<0.001	yes	no
Correctness	12	<<0.001	yes	<=0.003	yes	no
Confidence	12	<<0.001	yes	<=0.002	yes	no
Speed	12	<<0.001	yes	<<0.001	yes	no

Table 32 Results by Question

Correctness for Q5 Vitamin, Q5 JIMI, Q7 Vitamin, and Q11 VITAMIN are constant for the CBD. Additionally, for the washout test, Q7 VITAMIN and Q12 VITAMIN are constant. For these questions and treatments, all subjects answered the task correctly.

Once again, the enhanced interface system displays significant results for questions of more complex task-types. None of the task-type I questions has significant results in all

three factors. One of the task-type II questions (Q7) has significant results for all three factors, and two of the type III task questions (Q11 and Q12) have significant results for all three hypothesis factors.

The minor differences between the CBD and the washout tests appear to be errors rather than actual significant differences. The only three deltas were all in subjective *confidence* scores with one question indicating more confidence in VITAMIN and two indicating more confidence in JIMI. The washout assumption holds for by-question analysis. Additionally, Table 32 also reveals the "expert effect" on Question 10 that was first identified by the qualitative analysis. Question 10 is the only question that fails to reject the null hypothesis for *speed*.

The results have been presented and analyzed. Chapter Five will now explore the meaning of these analyses and suggest directions and implications for future research.

CHAPTER 5: CONCLUSION AND FUTURE WORK

> Perfection of means and confusion of goals seem- in my opinion- to characterize our age. – Einstein [EIN50]

A review of the usability body of knowledge in chapter two revealed several opportunities for original contributions. Among these opportunities was the need for additional experimentation in computer science, particularly usability research with novel systems, tasks, and environments. The literature stressed the need for usability studies "in context" whereby "a specified set of users can achieve a specified set of tasks in a particular environment," according to the ISO usability standard [ISO91]. Additional researchers have shown empirically that usability enhancements that are effective in some environments are not necessarily effective in all specific, well-defined contexts. The specific usability context of this research for which no empirical results were reported is in the environment of remote usability evaluation of interface systems to agent-mediated distributed heterogeneous legacy data sources. No previous experiments addressed the research question.

To simplify discussion of these issues in this context, two new expressions were defined for this research. This need for compact terms resulted in the development of the theoretical concept of interface system *coherence* and interaction metaphor of *indirect manipulation*. An additional opportunity in remote usability evaluation resulted in the specification of a novel system with which to conduct remote testing. Chapter Three grounded these theoretical ideas with the development of two prototype indirect manipulation interface systems (IMIS), VITAMIN and JIMI, and a remote usability evaluation system, JUICE. VITAMIN was developed with enhanced coherence in comparison to the more traditional JIMI system. The challenging domain of Army War Reserve equipment readiness was selected to provide a concrete comparison of the enhanced VITAMIN system for agent-mediated heterogeneous legacy data sources. The results of the experiments were analyzed in chapter four. This analysis revealed that VITAMIN outperformed JIMI by every measure in a statistically significant manner.

This final chapter of the dissertation explores the meaning of this research and these analyses and then explores suggestions and implications for future research. First, the experimental conclusions are addressed along with their practical applications and contributions to the body of knowledge. Finally, suggestions for future work and an overall summary of the dissertation are reported.

5.1 Conclusions and Contributions

Statistically significant results were obtained from the remote evaluations for all four indicators of coherence, with VITAMIN providing an improvement over JIMI as measured by: (1) the number of *correctly* answered tasks; (2) the user *confidence* in correctly answered tasks; (3) the time to answer tasks correctly (*speed*); and (4) the user *satisfaction* with the interface. These results held true when all of the cases were considered and when the more conservative test that discarded the second treatment (the washout assumption test) was calculated. These results provide unique empirical support by remote usability evaluations for the use of enhanced indirect manipulation interface systems to improve the coherence of agent-mediated legacy data.

Thus, the research question, listed below, may be answered in the affirmative:

Can an enhanced indirect manipulation interface system add coherence to agent-mediated legacy data for users performing representative tasks?

This question is important because it addresses several open issues identified in Chapter Two, namely the need for remote evaluation, the need for usability testing in a specific, well-defined context, and the increasingly common requirement to provide interface systems which enable efficient access to complex heterogeneous data. The next two paragraphs provide a review of these issues and then detailed conclusions are examined.

According to Plaice [PLA95] and Tichy [TIC98], remote usability testing and experimentation is vitally important to the field of computer science. This dissertation relies on usability theories to construct interface and evaluation systems, and the research uses an agent-based approach to access the legacy data sources. According to Wegner and Doyle [WEG96], usability and agents are extremely active and important research

areas. Additionally, the research community agrees that usability cannot be separated from the overall context. This context includes the users, the tasks, and the overall organizational and environmental situation. Tullis and Kodimer [TUL95] showed experimentally that a command line system surpassed enhanced interface systems under certain circumstances. Clearly, enhanced techniques in one context do not always provide improvements in all additional environments.

Particularly in homogeneous environments, interface usability in general has been well researched according to Jordan [JOR98]. However, many diverse organizations rely on information systems that depend heavily on distributed heterogeneous legacy data sources. These legacy data sources introduce a number of significant problems when they must be aggregated and efficiently displayed to remote users. Nielsen [NIE94] and Tullis [TUL93] have investigated these vital issues. Nevertheless, Shneiderman [SHN97a] describes remote interface systems to complex distributed data as a topic requiring further investigation, and empirical studies have not been published that examine such systems.

This dissertation investigated this topic by designing two indirect manipulation interface systems (IMIS). Both of these systems enable access to the complex distributed heterogeneous data sources through the HERMES mediator agent. The VITAMIN system was designed as an enhanced methodology interface system with visual cues grounded in usability theory. JIMI was designed to emulate a traditional interface system that would serve as the control treatment. The novelty of this approach is not simply in the comparison of the application of visual cues to interface systems, but instead, is in their application for an empirical evaluation of an indirect manipulation interface system in a particular context.

In general, previous research has shown that enhanced interface systems offer advantages; several studies cited here support this finding for direct manipulation systems in homogeneous environments. However, these results depend a great deal on the tasks and the specific environments. Studies have also demonstrated that very simple interface systems may surpass enhanced systems in specific circumstances.

The literature has several examples of interface agent and remote direct manipulation systems that are similar to indirect manipulation systems. Nonetheless, these citations often simply demonstrate that such systems are possible and do not evaluate their overall efficacy in comparison to competing or complementary approaches. Very few experimental results have been reported in this regard; none address the research problem.

The practical significance of these results is that the indirect manipulation interface system that provided enhanced representation of the legacy data, VITAMIN, proved to be more *satisfactory* than the traditional JIMI system for the users to *correctly, confidently, and speedily* answer representative tasks.

The most dramatic impact was on the time to answer the tasks correctly, "speed." The mean user time for the subjects using the VITAMIN system was 34 seconds, with a standard deviation of 8 seconds. The mean time with the JIMI system was 82 seconds with a standard deviation of 21 seconds. The practical impact of these results is apparent if many queries are required to answer many tasks. Solving tasks with the JIMI system takes almost three times as long as solving the same tasks with the enhanced-coherence VITAMIN system. Speed was the only factor for which VITAMIN showed statistically significant results for all three task-types. In fact, the null hypothesis was rejected for every question except the "expert effect" on question 10. (The "expert effect" refers to several subjects who used a shortcut for question 10, a type III task, by issuing fewer carefully crafted queries instead of many single dimension queries required by the VITAMIN system.) Analysis of the speed factor provides extremely strong evidence that the theory-based enhanced visual cues of the VITAMIN system provided greater coherence and assistance for subjects performing representative tasks.

On the other hand, perhaps the most important variable is the number of queries answered correctly. For the given tasks, users who used VITAMIN answered approximately one more question correctly than users who used JIMI. Again, in a crucial situation, answering all 12 tasks correctly may be extremely important. Correctness is probably less sensitive than speed because, given enough time, the subjects were able to

arrive at the correct answer. If time to answer the tasks were held constant to an amount between the VITAMIN and JIMI mean times, the results suggest that far fewer tasks would have been answered correctly by subjects using the JIMI system.

The results by task-type also support the increased benefit of the VITAMIN system for tasks requiring more analysis according to Bloom's taxonomy. For the washout assumption test on the type I tasks, VITAMIN's statistically significant advantage was limited to the speed hypothesis. Task-type II showed advantages for speed and correctness, and task-type III demonstrated advantages for all three hypothesis factors. Similar results were presented for the by-question analysis. These results support the observation that the complex task-types gained more from the VITAMIN system's enhanced coherence.

Although there were statistically significant results for the *user confidence in correctly answered tasks*, the results for both IMIS were between "Very" and "Extremely" confident. The analysis revealed that VITAMIN was closer to "Extremely" and JIMI was closer to "Very." This suggests that when the subjects arrived at the correct answer, they were generally confident that they had done so. Yet, these aggregated overall results tend to hide stronger support for the VITAMIN system that is evident when the by-task and by-question results are analyzed. In the by-question results, one of the type I tasks (Q6) has a significant result. All type II tasks and all type III tasks except Q10 have significant results.

Finally, on every measure of satisfaction (Easy to use, Helpful, Perform Faster, Information Quality, Interface Quality, Learn more about the data, Enjoyable, and Useful), the subjects preferred the VITAMIN system to the JIMI system. For the overall satisfaction, subjects "Disagreed" that they were satisfied with JIMI and "Agreed" that they were satisfied with VITAMIN. Satisfaction was specified as the overall system satisfaction and was not measured by task type or by questions. These same trends held for each of the eight satisfaction factors as reported in Section 4.6 and shown in Tables 27 and 28.

The only two of these eight satisfaction factors for which JIMI was rated better than neutral were “The system was helpful” and “The system provided high interface quality.” On the other hand, all of the VITAMIN ratings were above neutral with most of them closest to the “Agree” rating. These satisfaction ratings were supported and elaborated by the qualitative comments. Comments about VITAMIN included the following: “VITAMIN would be much better for new workers,” “much easier to use,’ and “faster.” There were several positive comments about JIMI, but most of the comments about JIMI were negative. None of the subjects expressed any issues of fatigue, lack of motivation, or stress although one subject who did not complete the experiment expressed this frustration.

> The whole thing makes absolutely no sense. What do all those numbers mean? What are all those buttons I’m pushing? What is JAVE (sic), JUICE, JIMI, and all that? What is an interface? What is GUI? What is it? What does it do? What are those numbers and buttons?

Evidently, this subject did not understand anything about the experiment or the domain. Fortunately, the vast majority of subjects fared much better and completed the experiments.

All subjects used for this experiment were college freshman. These students had no background in logistics. Despite this lack of knowledge and experience, they were able to use the interface to answer nearly all of the representative tasks correctly. According to White [WHI98], the typical real-world user is a college-educated analyst with many years of experience in the subject domain. Additionally, three army officers and college professors took part in comparative experiments. The outcome of these experiments was very similar to that of the dissertation experiments presented here. The qualitative results of the dissertation experiments indicate that several users approached near expert level of understanding in this admittedly narrow portion of the domain. These few subjects expressed an appreciation for the ability to get directly at the data to solve complex tasks. These sentiments were expressed in qualitative comments that these subjects could “get

more information," "directly access the information," and "go directly to the data needed." This was demonstrated quantitatively as the "expert effect" on question 10.

The subjects who preferred JIMI also expressed concerns about the learning curve and the training value of JIMI. All subjects who commented on the ease of learning agreed that VITAMIN was "much easier to learn" and much better for new users. The quantitative results and qualitative comments also support common findings that interface systems that provide visual cues are easier to learn, understand, and use. Although common, this result is not universal for all users, tasks, and environments. This is another reason why this research with distributed agent-mediated heterogeneous legacy data sources is important. The consequences of this research may indicate that, with an enhanced interface system, expert analysts may not be required for routine tasks in this domain. Soldiers may be able to perform this analysis directly, perhaps even from networked laptops in a deployed field environment.

In addition to the JIMI and VITAMIN, a novel remote usability evaluation system was specified. This third system, the Java Usability Interface Comparison and Evaluation (JUICE) system, was developed as an innovative remote usability evaluation system to enable the determination of whether the enhanced coherence methodology in VITAMIN is superior to the traditional methodology represented by JIMI. JUICE proved to be an invaluable asset in enabling these experiments.

The implementations of JIMI, VITAMIN, and JUICE applied usability theories. Agent theories were also applied to access the legacy data; however, no original claims are made as to agent algorithms for this work. Nevertheless, a substantial additional effort was invested to solve the problems of decoding and accessing the distributed heterogeneous legacy data sources required for the test of indirect manipulation. As discussed in the *Methodology* chapter, an agent-based approach to accessing this data was selected. The challenging domain of Army War Reserve equipment readiness was chosen to provide a concrete comparison of the enhanced IMIS with the tradition methodology. The results of these experimental comparisons were analyzed to reveal

that the enhanced coherence methodology was statistically significantly superior to the traditional methodology by every hypothesized measure.

It is important to note that the existing usability literature described possible enhancements to interface systems in particular environments. No existing literature provided empirical analysis of usability in an agent-mediated legacy data environment with different levels of tasks and visual cues. The research presented here applied this grounded theory to the challenging domain and reports the empirical results in this dissertation. To simplify discussion of these issues in this context, two new expressions were defined for this research.

These two theoretical concepts are *coherence* and *indirect manipulation*. The concept of coherence was defined as shorthand for a concise usability metric for statistical analysis. Some limitations of this concept will be discussed in the next section on future work and research opportunities. Indirect manipulation was defined as an interaction metaphor between the current metaphors of direct manipulation and interface agents. Likewise, the interaction metaphor of indirect manipulation must be studied further before it will become a firm part of the usability lexicon. In any case, the concept of coherence and the interaction metaphor of direct manipulation allow this dissertation to describe these notions succinctly. Future research opportunities in these usability issues as well as issues of methodology are described in the next section.

5.2 Additional Research Opportunities

> If they don't depend on true evidence, scientists are no better than gossips. - Fitzgerald [FIT90]

This research offers many avenues for academic and practical additions and extensions. In this section, ideas for efforts that build on this work are offered in the areas of usability and methodology. Future work in usability includes improvements to the *indirect manipulation* interface systems, contributions to remote evaluations, and elucidation of the concept of *coherence*. Future work in methodology includes advancing

heterogeneous data access, expanding to additional functional domains, and improving the empirical design and analysis.

5.2.1 Future Work in Usability

Future researchers and practitioners may make use of and improve the interaction metaphor of *indirect manipulation.* Indirect manipulation is used here as a description of a compromise between the competing metaphors of direct manipulation and interface agents. Perhaps there is a better term such as "mediated manipulation" to describe this metaphor or perhaps future researchers will see erosion of the walls between direct manipulation and interface agents. Future researchers may classify interaction metaphors by contrasting traditional and innovative metaphors and providing a taxonomy for distinguishing them in several dimensions. Such research would aid future empirical research to clearly determine the boundaries and intersections of various interaction metaphors. The term "indirect manipulation" was inspired, in part, because it provides clear contrast with direct manipulation and because of its similar usage in the field of virtual reality.

Improvements could also be made to the indirect manipulation interface systems (IMIS), JIMI, and VITAMIN. New IMIS may be constructed that explore additional unique collections of enhancements. They may also explore additional interaction metaphors and additional usability factors in additional contexts to further specify the usability body of knowledge. Future practitioners may also implement IMIS, with a different interface system instead of Java, such as X, XML, ActiveX, or some other distributed application environment. These implementations may demonstrate that the IMIS is not dependent on a single underlying instantiation. Based on the qualitative comments and usability research, an important interface system design criteria is the enabling novice access without hindering advanced or expert users. Finally, the IMIS may be implemented with additional agent characteristics such as the ability to communicate using CORBA or KQML and the ability to locate data sources and mediator agents dynamically. These agentized IMIS may meliorate some of the response time issues that currently draw a sharp distinction between direct manipulation and indirect manipulation interface systems.

In addition to improvements to the IMIS, the interaction between the IMIS and the remote evaluation system, JUICE, may be improved. This is motivated by the following two comments expressed by the research subjects.

> On the VITAMIN system, a bigger viewer interface could allow the user to view more information and not have to open and close as much.

> I found it somewhat difficult to line up the results with the corresponding columns at the top, as I got deeper and deeper into a ship's information. It should perhaps be presented in a larger sized square with lines separating the columns to make it easier to read.

Several of the issues identified in these comments are caused by the need to limit the IMIS to approximately half of the screen in order to accommodate the representative tasks and other questions that JUICE enables. Possible methods to mitigate these issues are to redesign the JUICE system such that the IMIS display can be more representative of its typical size as a stand-alone system.

JUICE enables remote experimentation for two reasons. First, empirical verification is essential to the field of computer science according to Plaice [PLA95] and Tichy [TIC98]. Secondly, the "remoteness" of the ubiquitous networked environment is not easily duplicated in a usability laboratory. According to Hartson et al. [HAR96b], "the network itself and the remote work setting have become intrinsic parts of usage patterns, difficult to reproduce in a laboratory setting." Therefore, in many current settings, remote evaluation offers an opportunity for more realistic results than laboratory evaluation.

In this vein, JUICE may be used by future researchers to evaluate innovative interfaces and future improvements to JUICE may allow the investigation of additional usability factors. This may result in common set of factors and metaphors from the interaction taxonomy that may be generalized into more accurate interface system heuristics for

practitioners. JUICE may also be improved to be a more general web-based experimental enabler.

Nebesh [NEB97] and Samadi [SAM97] demonstrated that experiments could be successfully accomplished using the web. These experiments were not usability evaluations; however, they did compare various systems and record questionnaire and timing information. Inspired by these experiments, enhancements to JUICE could allow it to read the survey and task questions from arbitrary data sources and further, to allow the results to be stored in dynamically specified data sources. Presently, the task and survey questions are encoded directly in the JUICE class files. A more flexible approach will obtain the questions from a database.

In addition, a customizable Web-based perl CGI script to administer and conduct web page evaluations was developed by Perlman [PER98]. This script has no option to record task response times or present various systems for comparison; however, it allows dynamic specification of the web interface systems under evaluation. Such improvements to JUICE could enable researchers to apply remote evaluation techniques to a wider variety of systems. Currently, the JIMI and VITAMIN Java class files are encoded directly into JUICE. A more general approach will allow the researcher to dynamically specify the evaluated interface systems with a Uniform Resource Locator (URL) or alternate specification method.

Additional improvements to JUICE may include a modular design for easier maintenance. Enhancements in question access, interface specification, and modularity will allow JUICE to be used for additional tasks and interface systems. All of these improvements could be implemented in a straightforward manner by future researchers. These upgrades may be validated with further formative evaluations of JUICE. Finally, researchers may conduct additional surveys to locate similar academic or commercial evaluation systems that may become available in the future. These remote evaluation systems may be compared and contrasted with JUICE. At the time of this research, no remote evaluation system was identified that could enable the IMIS comparison and record the desired set of usability factors.

To describe this set of usability factors concisely, the concept of coherence was introduced as a usability metric for statistical analysis. These factors include specific tasks performed *correctly* (effectively), *confidently, speedily* (efficiently), *and satisfactorily*. All of these factors may be measured by remote evaluation systems. This broad characterization of coherence is established to denote these four usability factors without requiring a precise set of interface devices or metaphors.

Future work with the coherence metric may refine one or more of the coherence factors (*correctly, confidently, speedily, and satisfactorily* - C^2S^2). Additional revisions may be made to the satisfaction factors, possibly reducing the required number from eight to four in order to make them more concise. These revisions and refinements may be based theory or on additional validity and reliability studies. These revisions may provide empirical support for the concept of coherence and bring about a more useful metric for usability evaluations.

Additionally, the coherence metric described in this research may only be used to compare two interfaces. As currently defined, researchers may state that one interface system is more coherent than another interface system for a particular set of users and tasks in a specific environment. A future improvement may allow the coherence metric to label interface systems according to ratings on a predefined scale. Such an improvement may allow practitioners to state that a certain interface system has a coherence rating of “8.6”, for instance. To accomplish such a rating system, to move the coherence metric from a comparison to a rating, a research framework must be developed. This framework must account for the tasks, subjects, and environment as well as the essential data types, locations, and access times for the underlying data sources.

A multidimensional taxonomy may be developed such that future practitioners may be able to enable standard interface system enhancements that will increase the coherence of systems in a predictable way for a wide variety of tasks. As a starting point, the tasks may be placed on a task-type axis of the coherence taxonomy according to Bloom’s taxonomy. A second axis may consider interface system enhancements such as improved data selection instead of data entry, improved visual clarity with familiar color effects,

and improved representation with an aggregating tree structure. The third axis in a coherence taxonomy may consider the context and environment of the representative tasks under consideration.

Coherence is also defined here as a characteristic of interface systems that present interaction components in a logical and consistent manner. Future researchers and practitioners may make use of and refine this sense of the concept of coherence. In order to prove the existence and reliability of coherence, at least two additional usability metrics that provide the same result must be compared. Additional work may include defining two terms instead of overloading "coherence" as both a metric and an interface system characteristic. "Information Efficacy" is one alternative phrase that may be used instead of "coherence" as the characteristic term to describe the usability and usefulness of the information that a user derives from displayed data. The term "coherence" was selected by this research because of its analogous connotation in the field of computer graphics.

In addition to the future theoretical work in the usability field, several opportunities are available in the implementation and methodology. These suggestions are provide in the next section.

5.2.2 Future Work in Methodology

Future work in methodology may involve improving heterogeneous data access, investigating alternate functional domains, and improving the experimental design and analysis.

Future work in data access may include several techniques to eliminate the reliance on the statically specified HERMES mediator agent. These include allowing the IMIS to dynamically determine the mediator agent or to allow alternate data sources to be used. Furthermore, entirely different agent architectures may be implemented. These improvements will allow the IMIS to be applied to a wider class of problems that do not rely on a fixed agent to mediate data access. Another improvement may use a different interoperability enabler altogether, instead of agents. Perhaps object request brokers

(ORBs), other middleware, or extensible markup language (XML) may be used to directly apply the IMIS methodology to existing applications that use these technologies.

Furthermore, service improvements may also allow the set of solutions to be expanded. These improvements to the data access services include additional security and speed optimizations. Security may be improved with virtual private networks (VPN), secure socket layer (SSL), or other methods. A portion of this implementation would require the investigation of solutions to various security implications of the mediated data gathering and collaborative planning environment. Speed may be improved with data replication, duplication, enhanced quality-of-service (QOS) networking, and data caching. In addition to simply expanding the functional domains with existing data sources, future research may also develop a unified framework that encompasses multimedia and GIS data sources. Such a system may address the computer-assisted extraction of meaningful information and knowledge from large quantities of remote data, text documents, and other media. All of these techniques serve to expand the coverage of IMIS to provide coherent representations of various functional domains.

Future work in the experimental design and analysis may include using a different experimental design instead of a within subjects counterbalanced design (CBD). This may result in fewer cases per subject and would require a larger pool of subjects to arrive at the same number of cases for statistical analysis. Additional design changes include running the same experiment with remote subjects that are not constrained by the need to sign "Informed Consent" forms. In this way, the experiment may be opened to an entire Internet community. In terms of validity threats, the CBD is very strong at mitigating threats to internal validity; that is the ability to show that changes in the dependent variable are caused by changes in the independent variable. However, opening up the experiment to a wider variety of subjects may allow the results to be generalized to a wider population, which is the measure of external validity. Finally, this research anticipated a learning effect from the CBD. Although the washout assumption test did not reveal this effect, future work that employs a CBD must continue to test this assumption of independent samples.

Future advances to the data analysis may involve performing analysis on different variables from the four main factors identified for this research. Additional variables that are already collected by JUICE may reveal interesting correlations. These variables include total time, number of queries, time spent on incorrectly answered tasks and approximately seventy additional variables in the data at Appendix J.

Analysis of larger collections of these variables may require the employment of different analysis techniques such as multivariate analysis, the general linear model, and the calculation of correlations and factor analysis. These statistical techniques may result in a more rigorous treatment of the experimental results and discover interactions that have not been identified by the hypothesis testing described in this work.

This research was primarily concerned with several aspects of usability, but it also dealt with many additional aspects of the computer science body of knowledge from agents to security and information warfare. Any or all of these areas offer many opportunities for additional research and practical applications. The suggestions listed above represent the only a small number of ideas for future work that may be inspired by this research.

5.3 Dissertation Summary

This research presents the empirical results of remote usability experiments in a unique context. The principle benefit of this research is that it demonstrates that an enhanced indirect manipulation interface systems (IMIS) adds coherence to agent-mediated legacy data for users performing representative tasks in a specific, well-defined context. The analyses indicate strong support for the enhanced IMIS with statistically significant results obtained for all four indicators of coherence. The enhanced IMIS, VITAMIN, provided an improvement over the traditional IMIS, JIMI, for each coherence indicator: (1) the number of correctly answered tasks; (2) the user confidence in correctly answered tasks; (3) the time to answer tasks correctly; and (4) the user satisfaction with the interface.

To assist with this analysis, the theoretical concept of interface system *coherence* and the interaction metaphor of *indirect manipulation* were developed. The practicality of these

two concepts was demonstrated with a unique remote evaluation system, JUICE, which was also developed for this research.

In summary, several opportunities to contribute to the computer science body were identified from the literature review. These opportunities included empirical remote usability evaluations in unique and complex contexts, implementation of enhanced interface systems, definition of the terms *coherence* and *indirect manipulation*, and implementation of a remote usability evaluation system. These opportunities were each addressed by this dissertation research.

REFERENCE LIST

Works cited in the text of the dissertation.

[ADA96a] Adali, Sibel, Kasim Selcuk Candan, Yannis Papakonstantinou, and V. S. Subrahmanian. 1996. Query caching and optimization in distributed mediator systems. In *ACM SIGMOD International Conference on Management of Data*, 25: 137-148. New York: ACM Press.

[ADA96c] Adali, Sibel and V. S. Subrahmanian. 1996. Amalgamating Knowledge Bases, III - Algorithms, Data Structures, and Query Processing. *Journal of Logic Programming* 28, no. 2: 45-88. http://www.cs.umd.edu/projects/hermes/publications/abstracts/papers_akbiii.html

[AHD92] AHD. 1992. *The American Heritage(r) Dictionary of the English Language*. Electronic version licensed from INSO Corporation: Houghton Mifflin Company.

[ALD99] Alderson, Albert and Hanifa Shah. 1999. Viewpoints on Legacy Systems. *Communications of the ACM* 42, no. 3: 115-116.

[ARE94] Arens, Y., C. Y. Chee, C.-N. Hsu, H. In, and Craig A. Knoblock. 1994. Query Processing in an Information Mediator. In *ARPA/Rome Laboratory Knowledge-Based Planning and Scheduling Initiative Workshop*. Tucson, AZ. http://www.isi.edu/sims/papers/; homepage = http://www.isi.edu/sims/sims-homepage.html

[ARI98] Arisha, Khaled, Sarit Kraus, Fatma Ozcan, Robert Ross, and V. S. Subrahmanian. 1998. IMPACT: The Interactive Maryland Platform for Agents Collaborating Together. *IEEE Intelligent Systems* 14, no. 2: 64-72. http://www.cs.umd.edu/projects/impact/pubs.html

[ARN95] Arnold, Jesse C. and J. Susan Milton. 1995. *Introduction to Probability and Statistics : Principles and Applications for Engineering and the Computing Sciences*. McGraw-Hill Series in Probability.

[ASB98] ASB. 1998. *Interoperability*. Washington, DC: Army Science Board1998. Report. http://www.sarda.army.mil/SARD-ASB/acronyms.htm; http://www.sarda.army.mil/sard-zt/ASTMP98/astmp98.htm; http://www.sisostds.org/

[ASH98] Ashish, Naveen. 1998. Optimizing Information Mediators by Selectively Materializing Data. Doctoral Thesis Proposal, University of Southern California. http://www.isi.edu/sims/naveen/dc98.ps

[BEV91] Bevan, Nigel, Jurek Kirakowski, and Jonathan Maissel. 1991. What is Usability? In *Proceedings of the Fourth International Conference on Human-Computer Interaction*, 1: 651--655. Stuttgart. ftp://ftp.npl.co.uk/pub/hci/papers/usabilit.rtf

[BEV93] Bevan, Nigel and Miles Macleod. 1993. Usability assessment and measurement. In *The management and measurement of software quality*, ed. M Kelly: Ashgate Technical / Gower Press. ftp://ftp.npl.co.uk/pub/hci/papers/assess.rtf

[BLO56] Bloom, B.S. 1956. *Taxonomy of Educational Objectives: The Classification of Education Goals*. New York: David McKay.

[BOC93] Bock, Peter. 1993. *The Emergence of Artificial Cognition : An Introduction to Collective Learning*: World Scientific Pub Co.

[BON98a] Bond, William L. 1998. "Digitization: View of the STRICOM Commanding General." A paper delivered at the NDIA 27th Annual Government/Industry Conference, 25-Sep 1998. http://www.stricom.army.mil

[BON98c] Bontempo, Charles and George Zagelow. 1998. The IBM data warehouse architecture. *Communications of the ACM* 41, no. 9: 38-48. http://www.acm.org/pubs/citations/journals/cacm/1998-41-9/p38-bontempo/

[BOR94] Borenstein, Nathaniel S. 1994. *Programming As If People Mattered: Friendly Programs, Software Engineering, and Other Noble Delusions*. Princeton: Princeton University Press. http://www.amazon.com/exec/obidos/ASIN/0691037639

[BRA97] Bray, T., J. Paoli, and C. Sperberg-McQueen. 1997. Extensible Markup Language (XML). *The World Wide Web Journal* 2, no. 4: 29--66. http://www.w3.org/TR/1998/REC-xml-19980210

[BRO96] Brooke, John B. 1996. SUS: A Quick and Dirty Usability Scale. In *Usability Evaluation in Industry*, ed. P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland. London: Taylor & Francis. http://www.redhatch.co.uk/sus.html

[BUN97] Buneman, Peter, Louisa Raschid, and Jeffrey D. Ullman. 1997. Mediator languages -- a proposal for a standard. *SIGMOD Record* 26, no. 1: 39. http://www.acm.org/sigmod/record/issues/9703/index.html

[BUT74] Butler, Sir William. 1874. Its fighting done by fools and its thinking by cowards.

[BYR99] Byrnes, Ronald, Joseph H. Schafer, and John A. Marin. 1999. *Dynamic Scenario Builder*.:1999. Technical Report to the Simulation, Training, and Instrumentation Command (STRICOM), OAIAE-TR 1999 02 01.

[C4I98a] C4ISR-AWG. 1998. *C4ISR Architecture Working Group Final Report*. Washington, DC: Assistant Secretary of Defense for Command, Control, Communications and Intelligence, 18-Apr 1998. Working Group Report, C4ISR Architecture Working Group - Final Report. http://www.c3i.osd.mil/doc/index.html

[C4I98b] C4ISR-AWG. 1998. *Levels of Information Systems Interoperability (LISI)*. Washington, DC: Assistant Secretary of Defense for Command, Control, Communications and Intelligence, 30-Mar 1998. Working Group Report, C4ISR Architecture Working Group - LISI. http://www.c3i.osd.mil/AWG_Digital_Library/pdfdocs/lisi.pdf; http://www.cisa.osd.mil/organization/architectures/LISI/

[CAF98] Caftori, Netiva, Nathaniel S. Borenstein, and C. Hoelscher. 1998. "One planet, one net: Principles for the Internet era." A paper delivered at the World Conference on the World Wide Web, Orlando, 1998.

[CAM98a] Campbell, Lieutenant General William H. 14 May 1998. Questions for the Army Director of Information Systems for Command, Control, and Communications at West Point. Meeting, USMA D/EECS.

[CAN96f] Candan, Kasim Selcuk, V. S. Subrahmanian, and V. Rangan. 1996. Towards a Theory of Collaborative Multimedia. In *IEEE International Conference on Multimedia Computing and Systems*. Hiroshima, Japan. http://www.cs.umd.edu/projects/hermes/publications/abstracts/ttcm.html

[CAR83] Card, Stuart K., Thomas P Moran, and Allen Newell. 1983. *The psychology of human-computer interaction*. Hillsdale, N.J.: L. Erlbaum Associates.

[CHA97a] Chaudhuri, Surajit and Umeshwar Dayal. 1997. Data warehousing and OLAP for decision support. In *ACM SIGMOD International Conference on Management of Data*, ed. Joan Peckham: 507-508. Tucson, AZ. http://www.acm.org/pubs/contents/proceedings/mod/253260/index.html

[CHA97b] Chaudhuri, Surajit and Umeshwar Dayal. 1997. Data Warehousing and OLAP for Decision Support (Tutorial). In *International Conference on Management of Data*, ed. Joan Peckham: 507-508. Tucson, Arizona.

[CHE91] Chelimsky, Eleanor. 1991. *Designing Evaluations*. Washington, DC: US General Accounting Office.

[CHI87] Chin, John P., Virginia A. Diehl, and Kent L. Norman. 1987. *Development of an instrument measuring user satisfaction of the human-computer interface*.: University of Maryland, College Park, September 1987. Technical Report, CS-TR-1926.

[CHI88] Chin, John P., Virginia A. Diehl, and Kent L. Norman. 1988. Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface. In *Proceedings of ACM CHI'88 Conference on Human Factors in Computing Systems*: 213--218.

[CLA97] Clark-Carter, David. 1997. *Doing Quantitative Psychological Research : From Design to Report*. East Sussex, UK: Psychology Press.

[COD70] Codd, E. F. 1970. A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM* 13, no. 6: 377-387.

[COL00] Colarusso, Michael. March 14 2000. Historical Relationship of Ease of Force Projection to Use of Force. Personal Email.

[COL42] Collingwood, R. G. 1942. *The New Leviathan*. Quoted in The Columbia Dictionary of Quotations: Columbia University Press.

[COL95] Colonna-Romano, John and Patricia Srite. 1995. *The Middleware Source Book*: Digital Press. http://www.amazon.com/exec/obidos/ISBN=1555581358

[COL96] Colliat, G. 1996. OLAP, Relational, and Multidimensional Database Systems. *SIGMOD Record* 25, no. 3: 64. http://www.acm.org/sigmod/record/issues/9609/index.html

[CON89] Conran, Sir Terence. 1989. Perhaps believing in good design is like believing in God, it makes you an optimist. *Daily Telegraph. Quoted in The Columbia Dictionary of Quotations*, 12 June.

[CRO97] Cross, Paul. 1997. "Architecting Connectivity Solutions Using Two And Three Tier Architectures." A paper delivered at the Europe, Middle East & Africa Oracle User Group Conference (EOUG97), 1997. http://www.fors.com/eoug97/papers/0304.htm

[CUV99] Cuviello, Peter. 13 Apr 1999. Chief of Signal Discussion with Information Operations Officers at West Point. Meeting, USMA D/EECS. http://www.gordon.army.mil/cmdgrp/cg.htm

[DAL97] Dale, Jonathan. 1997. A Mobile Agent Architecture for Distributed Information Management, University of Southampton. http://www.mmrg.ecs.soton.ac.uk/publications/archive/dale1997a/html/papers/thesis.htm

[DAR98] DARPA. 1998. *I3 - Intelligent Integration of Information.* 1998. Web Site. http://mole.dc.isx.com/I3/html/briefs/I3brief-Long/index.htm; http://dc.isx.com/I3/

[DAV83] Davis, Randall and Reid Smith. 1983. Negotiation as a Metaphor for Distributed Problem Solving. *Artificial Intelligence* 20, no. 1: 63--109.

[DAV89] Davis, Fred. D., R. P. Bagozzi, and P. R. Warshaw. 1989. User acceptance of computer technology: A comparison of two theoretical models. *Management Science* 35, no. 8: 982-1003.

[DAV93] Davis, Fred D. 1993. User Acceptance of Information Technology: System Characteristics, User Perceptions and Behavioral Impacts. *International Journal of Man-Machine Studies* 38, no. 3: 475--487.

[DAV95] Davis, F. D. and V. Venkatesh. 1995. Measuring User Acceptance of Emerging Information Technologies: An Assessment of Possible Method Biases. In *Proceedings of the 28th Annual Hawaii International Conference on System Sciences. Volume 4: Information Systems - Collaboration Systems and Technology Organizational Systems and Technology*, ed. Jay F. Nunamaker and Ralph H. Sprague: 729--736. Los Alamitos, CA, USA: IEEE Computer Society Press.

[DEC97a] Decker, Keith S., Katia P. Sycara, and Mike Williamson. 1997. Middle-Agents for the Internet. In *15th International Joint Conference on Artificial Intelligence*: 578-583. San Francisco: Morgan Kaufmann Publishers. http://www.cis.udel.edu/ decker/pubs/middleagents.ps

[DEC97b] Decker, Keith, Anandeep Pannu, Katia Sycara, and Mike Williamson. 1997. Designing Behaviors for Information Agents. In *Proceedings of the First International Conference on Autonomous Agents (Agents'97)*, ed. W. Lewis Johnson and Barbara Hayes-Roth: 404--412. New York: ACM Press. http://www.cis.udel.edu/ decker/pubs/aa97-final.ps

[DEP00] Deponai, Jerry. March 15 2000. Moral Implications of Research. Personal Email.

[DIR77] Dirac, Paul Adrien Maurice. 1977. in science one tries to tell people. In *Mathematical Circles Adieu*, ed. H. Eves. Boston: Prindle, Weber and Schmidt.

[DIS96a] DISA. 1996. *Technical Architecture Framework for Information Management Version 3.0*. Washington, DC, 30-Apr 1996. Report, TAFIM V3. http://www-library.itsi.disa.mil/tafim/

[DIS96b] DISA. 1996. *Volume 1, Overview , Technical Architecture Framework for Information Management Version 3.0*. Washington, DC, 30-Apr 1996. Report, TAFIM V3. http://www-library.itsi.disa.mil/tafim/

[DIS96c] DISA. 1996. *Volume 3 , Architecture Concepts and Design Guidance, Technical Architecture Framework for Information Management Version 3.0*. Washington, DC, 30-Apr 1996. Report, TAFIM V3. http://www-library.itsi.disa.mil/tafim/

[DIS98a] DISA. 1998. *Data Standardization Procedures*. Washington, DC: Assistant Secretary of Defense for Command, Control, Communications and Intelligence, 2-Apr 1998. Report, DoD 8320.1-M-1. http://web7.whs.osd.mil/html/83201m1.htm

[DIS98b] DISA. 1998. *DoD Joint Technical Architecture (JTA) Version 2.0*. Washington, DC: Defense Information Systems Agency, 26-May 1998. Report, JTA V2. http://www-jta.itsi.disa.mil/jta/jtav2_dnld.html; http://www-jta.itsi.disa.mil/; http://dii-sw.ncr.disa.mil/coe/

[DMS94] DMSO. 1994. *DoD Modeling and Simulation (M&S) Management*. Washington, DC: Defense Modeling and Simulation Office, 4-Jan 1994. Report, DoD Directive 5000.59. http://www.dmso.mil/dmso/docslib/mspolicy/directive.html

[DMS95b] DMSO. 1995. *DoD Modeling and Simulation (M&S) Master Plan*. Washington, DC: Defense Modeling and Simulation Office, October 1995. Report, DoD 5000.59-P. http://www.dmso.mil/dmso/docslib/mspolicy/msmp/1095msmp/

[DMS97] DMSO. 1997. *High Level Architecture (HLA) Glossary*. Washington, DC: Defense Modeling and Simulation Office, 3-Jun 1997. Report. http://hla.dmso.mil/hla/general/hlagloss.html

[DMS98] DMSO. 1998. *Glossary Of Modeling And Simulation (M&S) Terms*. Washington, DC: Defense Modeling and Simulation Office, 15-Jan 1998. Report. http://www.dmso.mil/dmso/docslib/mspolicy/glossary/

[DOD93a] DOD. 1993. *Defense Standardization Program (DSP) Policies & Procedures*.:, 1-Jul 1993. Report, DoD 4120.3-M. http://tecnet0.jcte.jcs.mil:9000/htdocs/teinfo/software/man1.html; http://www.nslc.fmso.navy.mil/techlog/nsg/dod4120.htm

[DOD93b] DOD. 1993. *Standard Simulator Data Base (SSDB) Interchange Format (SIF)*. Washington, DC: Department Of Defense, 17 June, with change pages and additions from MIL-STD-1821 NOTICE 1, 17 April 1994 included 1993. Report, MIL-STD-1821. http://www.vsl.ist.ucf.edu/groups/vtb/pages/sifDocForward.html

[DOD96] DOD. 1996. *Defense Acquisition Management Policies and Procedures*.:1996. Report, DODI 5000.2.

[DUN98] Duncan, James. 1998. *The Transportable Object Model for PERL*. 18-Jun 1998. http://www.hawk.igs.net/ jduncan/Tom/

[DYS95] Dyson, Freeman. 1995. The Scientist as Rebel. In *Nature's Imagination The Frontiers of Scientific Vision*, ed. J. Cornwell. Oxford: Oxford University Press.

[EIN50] Einstein, Albert. 1950. *Out of My Later Years*. Quoted in The Columbia Dictionary of Quotations: Columbia University Press.

[EIS89] Eisenhardt, K. 1989. Agency Theory: An Assessment and Review. *Academy of Management Review* 14, no. 1: 57-74. http://www.aom.pace.edu/

[ERI90] Erickson, Thomas D. 1990. Working with Interface Metaphors. In *Art of Human-Computer Interface Design*, ed. Brenda Laurel. New York: Addison-Wesley Pub Co. http://www.amazon.com/exec/obidos/ASIN/0201517973

[FIP91] FIPS. 1991. *Guideline: American National Dictionary for Information Systems*. Washington, DC: Federal Information Processing Standards Publications FIPS PUBS (NIST), 1-Feb 1991. Report, FIPS 11-3 (withdrawn).

[FIT90] Fitzgerald, Penelope. 1990. *The Gate of Angels (Herbert Flowerdew to Fred Fairly)*. Quoted in The Columbia Dictionary of Quotations: Columbia University Press.

[FRA97b] Franklin, Stan and Art Graesser. 1997. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In *Third International Workshop on Agent Theories, Architectures, and Languages (1996), published as Intelligent Agents III*, ed. Michael J. Wooldridge and Nicholas R. Jennings Jörg P. Müller, 1193: 21-36. New York: Springer-Verlag. http://www.msci.memphis.edu/ franklin/AgentProg.html

[FRI98a] Friedland, Liam. 1998. Accessing the data warehouse: designing tools to facilitate business understanding. *ACM interactions* 5, no. 1: 25-36. http://www.acm.org/pubs/citations/journals/interactions/1998-5-1/p25-friedland/

[FRO97] Frost, Rob. 1997. *Java(tm) Agent Template (JAT) Lite*. Stanford, CA: Stanford Center for Design Research and Enterprise Integration Technologies, Inc.1997. Report. http://cdr.stanford.edu/ABE/JavaAgent.html; http://cdr.stanford.edu/ProcessLink/ABE/; http://java.stanford.edu/

[GAR98] Gardner, Stephen R. 1998. Building the data warehouse. *Communications of the ACM* 41, no. 9: 52-60. http://www.acm.org/pubs/citations/journals/cacm/1998-41-9/p52-gardner/

[GAR99] Garson, G. Dave. 1999. *Statnotes: An Online Textbook. A Guide to Writing Empirical Papers, Theses, and Dissertations. A book under contract to Marcel Dekker, Publishers, for publication in 2001.* North Carolina State University, 1999. Web Site. 1999. http://www2.chass.ncsu.edu/garson/pa765/statnote.htm

[GIB99] Gibbins, Nick. 1999. *HCI and Agents.* 22-April 1999. Web Site. accessed May 1999. http://www.ecs.soton.ac.uk/ nmg97r/hci/agents/

[GOH94] Goh, Cheng Hian, S E Madnick, and M D Siegel. 1994. Context Interchange: Overcoming the Challenges of Large-Scale Interoperable Database Systems in a Dynamic Environment. In *3rd International Conference on Information and Knowledge Management*, ed. Nabil R Adam, Bharat K Bhargava, and Yelena Yesha: 337-346. Gaithersburg, MD. ftp://rombutan.mit.edu/pub/papers/cikm.ps

[GOL94] Goldstein, Jade and Steven F. Roth. 1994. Using aggregation and dynamic queries for exploring large data sets. In *ACM Conference on Human Factors in Computing Systems Celebrating Interdependence*: 23-29. Boston.

[GRA97] Granger, Mary J., Joyce Currie Little, Elizabeth S. Adams, Christina Björkman, Donald Gotterbarn, Diana D. Juettner, C. Dianne Martin, and Frank H. Young. 1997. Using information technology to integrate social and ethical issues into the computer science and information systems curriculum (report of the ITiCSE '97 working group on social and ethical issue in computing curricula). In *Integrating technology into computer science education: working group reports and supplemental proceedings*: 38-50. http://www.acm.org/pubs/citations/proceedings/cse/266057/p38-granger/

[GRO92] Gröller, Eduard. 1992. Coherence in Computer Graphics. Doctoral Thesis, Vienna University of Technology, The Institute of Computer Graphics. http://www.cg.tuwien.ac.at/research/theses/

[GRO95] Gröller, Eduard and Werner Purgathofer. 1995. *Coherence in Computer Graphics.*: Institute for Computer Graphics, Technical University Vienna, 1995 1995. Research paper, TR-186-2-95-04. http://visinfo.zib.de/EVlib/Show?EVL-1995-9; ftp://ftp.cg.tuwien.ac.at/pub/TR/95/TR-186-2-95-04Paper.ps.gz

[HAR96a] Hartley, Dawn. 1996. *Defense Information Infrastructure (DII) Shared Data Environment (SHADE) Capstone Document*. Falls Church, VA: Defense Information Systems Agency, 11-Jul 1996. Report, SHADE V1. http://sw-eng.falls-church.va.us/coe/docs/shade-capstone/

[HAR96b] Hartson, H. Rex, Jose C. Castillo, John Kelsa, and Wayne C. Neale. 1996. Remote Evaluation: The Network as an Extension of the Usability Laboratory. In *ACM Conference on Human Factors in Computing Systems: Common Ground*, ed. Michael J. Tauber, Victoria Bellotti, Robin Jeffries, Jock D. Mackinlay, and Jakob Nielsen, 1: 228-235. New York: ACM Press. http://www.acm.org/sigchi/chi96/proceedings/papers/Hartson/hrh_txt.htm

[HAR99] Harold, Elliotte Rusty. 1999. *XML Bible*. Chicago: IDG Books Worldwide. http://www.amazon.com/exec/obidos/ASIN/0764532367

[HAS99] Hassall, Jack, John Eaton, Gary Gray, Mike Wilcock, Arne-Jørgen Berre, and Tor Neple. 1999. *Applying ISO RM-ODP and UML in the Specification of Standard Interfaces to General Ledger Systems.* COMPASS: COMPonent based Accounting Service and System, 1999. Web Site. 1999. http://www.compassgl.org/newhometest/oops.html

[HCI99] HCIL. 1999. *QUIS: The Questionnaire for User Interaction Satisfaction.* Human Computer Interaction Laboratory, University of Maryland, College Park, 1999. Web Site. 1999. http://www.cs.umd.edu/hcil/quis/

[HEL95] Heller, Rachelle S. and C. Dianne Martin. 1995. A media taxonomy. *IEEE Multimedia* 2, no. 4: 36 - 45. http://computer.org/multimedia/mu1995/u4036abs.htm

[HFR93] HFRG. 1993. *The MUSiC Project (Measuring Usability of Software in Context).* Human Factors Research Group, 1993. Web Site. 1999. http://www.ucc.ie/hfrg/projects/music/index.html

[HFR99] HFRG. 1999. *Software Usability Measurement Inventory (SUMI).* Human Factors Research Group, 1999. Web Site. 1999. http://www.ucc.ie/hfrg/questionnaires/sumi/index.html

[HIG88] Higgins, George V. 1988. Data is what distinguishes the dilettante from the artist. *Guardian. Quoted in The Columbia Dictionary of Quotations*, 17 June 1988.

[HOW98] Howe, Denis. 1998. *The Free On-line Dictionary of Computing.* 1998. Web Site. http://wombat.doc.ic.ac.uk/foldoc/index.html; http://www.instantweb.com/foldoc/

[HQD92] HQDA. 1992. *Field Manual 7-20, The Infantry Battalion*. Washington, DC: Headquarters, Department of the Army. http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/7-20/toc.htm

[HQD94] HQDA. 1994. *Field Manual 100-1, The Army*. Washington, DC: Headquarters, Department of the Army. http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/100-1/toc.htm

[HQD96a] HQDA. 1996. *Field Manual 71-3, The Armored and Mechanized Infantry Brigade*. Washington, DC: Headquarters, Department of the Army. http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/71-3/toc.htm

[HQD96b] HQDA. 1996. *Field Manual 71-100, Division Operations*. Washington, DC: Headquarters, Department of the Army. http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/71-100/toc.htm

[HQD96c] HQDA. 1996. *Field Manual 100-6, Information Operations*. Washington, DC: Headquarters, Department of the Army. http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/100-6/toc.htm

[HQD96d] HQDA. 1996. *Field Manual 100-15, Corps Operations*. Washington, DC: Headquarters, Department of the Army. http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/100-15/toc.htm

[HQD96e] HQDA. 1996. *Field Manual 100-17-1, Army Pre-Positioned Afloat Operations*. Washington, DC: Headquarters, Department of the Army. http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/100-17-1/toc.htm

[HQD97a] HQDA. 1997. *Army Regulation 220-1, Unit Status Reporting for Field Organizations*. Washington, DC: Headquarters, Department of the Army.

[HQD98a] HQDA. 1998. *Army Regulation 380-19, Information Systems Security*. Washington, DC: Headquarters, Department of the Army.

[HUF95] Huff, Chuck and C. Dianne Martin. 1995. Computing Consequences: A Framework for Teaching Ethical Computing. *Communications of the ACM* 38, no. 12: 75-84. http://www.acm.org/pubs/toc/Abstracts/0001-0782/219687.html

[HUH98a] Huhns, Michael N. and Munindar P. Singh. 1998. Agents and Multiagent Systems: Themes, Approaches, and Challenges. In *Readings in Agents*. San Francisco: Morgan Kaufmann Publishers. http://www.amazon.com/exec/obidos/ISBN=1558604952

[HUH98b] Huhns, Michael N. and Munindar P. Singh, eds. 1998. *Readings in Agents*. San Francisco: Morgan Kaufmann Publishers. http://www.amazon.com/exec/obidos/ISBN=1558604952

[HUH98c] Huhns, Michael N. 1998. *Software Agent Technology: Assessment and Prospects*.: University of South Carolina1998. PowerPoint Presentation, CTI_AgentSurvey.ppt. http://www.ece.sc.edu/faculty/Huhns/

[HUH99] Huhns, Michael N. 1999. *Logistics Information Access via Cooperating Agents.* Center for Information Technology, University of South Carolina, October 1998 1999. Web Site. accessed May 1999. http://www.engr.sc.edu/research/cit/projects/ALP.html

[HUY98a] Hu, Yuh-Jong. 1998. Autonomous Security Agents: Negotiating Compatible Crypto Protocols on Behalf of the End-User. In *RSA Data Security Conference*. San Francisco. http://www.cs.nccu.edu.tw/ jong/pub/pub.html

[IEE90] IEEE. 1990. *IEEE Standard Glossary of Software Engineering Terminology*.:1990. Report, IEEE STD 610.12. http://standards.ieee.org/

[ISO91] ISO. 1991. *Ergonomic requirements for office work with visual display terminals*.: International Standards Organization1991, ISO 9241.

[ITM96] ITMRA. 1996. *Information Technology Management Reform Act (now the Clinger/Cohen Act)*. Washington, DC: U.S. Code1996. Report, Public Law 104-106. http://www.c3i.osd.mil/cio/references/itmra/itmra.Annot.html

[JAC96] Jaccard, James and Choi K. Wan. 1996. *LISREL Approaches to Interaction Effects in Multiple Regression*. Quantitative Applications in the Social Sciences. Thousand Oaks, CA: Sage Publications. http://www.amazon.com/exec/obidos/ASIN/0803971796; http://www.albany.edu/psy/fac_jjj.html

[JAI99] Jain, Anuj K., IV Aparico Manuel, and Munindar P. Singh. 1999. Agents for Process Coherence in Virtual Enterprises. *Communications of the ACM* 42, no. 3: 62--69. http://www.acm.org/pubs/articles/journals/cacm/1999-42-3/p62-jain/p62-jain.pdf

[JCS98a] JCS. 1998. *DoD Dictionary of Military and Associated Terms*. Washington, DC: Joint Staff, as amended through December 7 1998. Report, JP1-02. http://www.dtic.mil/doctrine/jel/doddict/

[JCS98b] JCS. 1998. *Joint Doctrine for Information Operations*. Washington, DC: Joint Staff, 9-Oct 1998. Report, Joint Pub 3-13. http://www.dtic.mil/doctrine/jel

[JCS98c] JCS. 1998. *Joint Vision 2010*. Washington, DC: Joint Staff1998. Report, JV2010. http://www.dtic.mil/doctrine/jel

[JOH95] Johnston, William E. and Deborah Agarwal. 1995. *The Virtual Laboratory: Using Networks to Enable Widely Distributed Collaboratory Science*. Berkeley: Ernest Orlando Lawrence Berkeley National Laboratory, 17 June 1995. white paper, NSF "vBNS and Networking and Application Researchers" workshop. http://www-itg.lbl.gov/~johnston/Virtual.Labs.html

[JOH96b] John, Bonnie E. and David E. Kieras. 1996. Using GOMS for user interface design and evaluation: which technique? *ACM Transactions on Computer-Human Interaction* 3, no. 4: 287-319. http://www.acm.org/pubs/citations/journals/tochi/1996-3-4/p287-john/

[JOR98] Jordan, Patrick W. 1998. *An Introduction to Usability*. London: Taylor & Francis. http://www.amazon.com/exec/obidos/ASIN/0748407626

[KAY90] Kay, Alan. 1990. User Interface: A Personal View. In *Art of Human-Computer Interface Design*, ed. Brenda Laurel. New York: Addison-Wesley Pub Co. http://www.amazon.com/exec/obidos/ASIN/0201517973

[KEI94] Keim, D. A., H.-P. Kriegel, and T. Seidl. 1994. Supporting Data Mining of Large Databases by Visual Feedback Queries. In *10th International Conference on Data Engineering*, ed. Ahmed K. Elmagarmid and Erich Neuhold: 302-313. Houston, TX: IEEE Computer Society Press.

[KER97] Kermarrec, Anne-Marie, Martin van Steen, and Andrew S. Tanenbaum. 1997. Coherence Models for Web Object-Based Systems. In *4th CaberNet Radicals Workshop*. Rethimnon, Crete: Vrije University. http://www.laas.research.ec.org/cabernet/research/radicals/1997/papers/coherence-kermarrec.ps

[KER98a] Kern, Lieutenant General Paul J. 7 Nov 1998. Questions for the Military Deputy to the Assistant Secretary of the Army (Research, Development, and Acquisition) at West Point. Meeting, USMA D/EECS. http://www.sarda.army.mil/bio/Kern_P_LTG.htm

[KIR93] Kirakowski, J. and M Corbett. 1993. SUMI: The Software Usability Measurement Inventory. *British Journal of Educational Technology* 24, no. 3: 210-212.

[KIR99] Kirakowski, Jurek. 1999. *Questionnaires in Usability Engineering: A List of Frequently Asked Questions. (2nd Ed.)*. Human Factors Research Group, Cork, Ireland., 1999. Web Site. http://www.ucc.ie/hfrg/resources/qfaq1.html

[KQM98] KQML. 1998. *Knowledge Query and Manipulation Language.* 1998. Web Site. http://kqml.org/tkqml

[KRI95] Kristof, Ray and Amy Satran. 1995. *Interactivity by Design: creating & communicating with new media*. Mountain View, CA: Adobe Systems Incorporated.

[KSE98] KSE. 1998. *Knowledge Sharing Effort.* 1998. Web Site. http://www.cs.umbc.edu/kse/

[KUM99] Kumara, Saundar R. T. 1999. *Decision Making in Logistics.* Alexandria, VA: Intelligent Design and Diagnostic Research Laboratory, Pennsylvania State University, 26-May 1999. Technical Report. http://www.iddr.ie.psu.edu/ dial

[LAB67] Labovitz, S. 1967. Some observations on measurement and statistics. *Social Forces* 46: 151-160.

[LAP99] LAP. 1999. *QUIS: The Questionnaire for User Interaction Satisfaction: About the QUIS, version 7.0.* Laboratory for Automation Psychology, University of Maryland, College Park, 1999. Web Site. 1999. http://www.lap.umd.edu/QUIS/

[LAR98] Larson, K. and M. P. Czerwinski. 1998. "Web Page Design: Implications of Memory, Structure and Scent for Information Retrieval." A paper delivered at the CHI ’98, Human Factors in Computing Systems, Los Angeles, April 21-23 1998. http://research.microsoft.com/research/ui/

[LAW98] Lawrence, S. and C. Giles. 1998. Searching the World Wide Web. *Science* 280, no. 5360: 98-100.

[LEW92] Lewis, James R. 1992. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. In *Proceedings of the Human Factors Society 36th Annual Meeting*, 2: 1259--1263.

[LEW95] Lewis, James R. 1995. IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction* 7, no. 1: 57--78.

[LIM96] Lim, Kai H., Izak Benbasat, and Peter A. Todd. 1996. An experimental investigation of the interactive effects of interface style, instructions, and task familiarity on user performance. *ACM Transactions on Computer-Human Interaction* 3, no. 1: Jan-37. http://www.acm.org/pubs/citations/journals/tochi/1996-3-1/p1-lim/

[LUK96] Luke, Sean, Lee Spector, and David Rager. 1996. Ontology-Based Knowledge Discovery on the World-Wide Web. In *AAAI Workshop on Internet-based Information Systems.*

[LUK97b] Luke, Sean, Lee Spector, David Rager, and James A. Hendler. 1997. Ontology-based Web Agents. In *First International Conference on Autonomous Agents*.

[MAC95] Mackinlay, Jock D., Ramana Rao, and Stuart K. Card. 1995. An Organic User Interface for Searching Citation Links. In *ACM Conference on Human Factors in Computing Systems*, 1: 67-73. http://www.acm.org/sigchi/chi95/proceedings/papers/jdm_bdy.htm

[MAE94a] Maes, Pattie. 1994. Agents that reduce work and information overload. *Communications of the ACM* 37, no. 7. http://www.ai.univie.ac.at/~paolo/lva/vu-sa/html/CACM-94/

[MAR85] Martin, James. 1985. *Fourth Generation Languages*. Englewood Cliffs, NJ: Prentice-Hall.

[MAR87] Margono, Sepeedeh and Ben Shneiderman. 1987. *A Study of File Manipulation by Novices Using Commands vs. Direct Manipulation*.: University of Maryland, College Park, June 1987. Technical Report, CS-TR-1775. http://www.cs.umd.edu/TRs/groups/HCIL.html

[MAR97] Martin, C. Dianne. 1997. The case for integrating ethical and social impact into the computer science curriculum. In *Integrating technology into computer science education: working group reports and supplemental proceedings*: 114-120. http://www.acm.org/pubs/citations/proceedings/cse/266057/p114-martin/

[MAR99b] Martin, David L., Adam J. Cheyer, and Douglas B. Moran. 1999. The Open Agent Architecture: A Framework for Building Distributed Software Systems. *Applied Artificial Intelligence* 13, no. 2-Jan: 91-128. http://www.ai.sri.com/pubs/papers/Mart:Open/document.ps.gz

[MAS89] Massie, Allan. 1989. *A Question of Loyalties (Colonel Fernie)*. Quoted in The Columbia Dictionary of Quotations: Columbia University Press.

[MIC97a] Microsoft. 1997. *Microsoft Bookshelf Computer and Internet Dictionary*. Seattle: Microsoft Corporation.

[MIC97b] Microsoft. 1997. *Microsoft Press Computer Dictionary*. Seattle: Microsoft Press.

[MIL56] Miller, G. A. 1956. The magical number seven plus of minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81-97.

[MIN85] Minsky, Marvin L. 1985. *The Society of Mind*. New York: Simon and Schuster.

[MIN96] Mine, Mark Raymond. 1996. Exploiting Proprioception in Virtual-Environment Interaction. Doctoral Dissertation, The University of North Carolina at Chapel Hill. http://www.cs.unc.edu/~mine/mine_dissertation.html

[MUR92] Murphy, Gregory L. and Douglas L. Medin. 1985. The Role of Theories in Conceptual Coherence. *Psychological Review* 92, no. 3: 289--315.

[MYE96a] Myers, Brad A., Jim Hollan, Isabel Cruz, Steve Bryson, Dick Bulterman, Tiziana Catarci, Wayne Citrin, Ephraim Glinert, Jonathan Grudin, and Yannis E. Ioannidis. 1996. Strategic directions in human-computer interaction. *ACM Computing Surveys* 28, no. 4: 794-809. http://www.acm.org/pubs/citations/journals/surveys/1996-28-4/p794-myers/

[NAG01] Nagy, Thomas J. 2001. "Ethical criteria for judging military research in the era of zero casualty warfare. Proposal in progress." A paper delivered at the Joint Services Conference on Professional Ethics, Washington, 2001. www.gwu.edu/~nagy/jscope2001.htm

[NAG99] Nagy, Thomas J. 1999. "Please refute the hypothesis: Core Values operate as a Smoke Screen to Evade the Law of Land Warfare." A paper delivered at the Joint Services Conference on Professional Ethics, Washington, January 28, 1999. www.gwu.edu/~nagy/jscope.ppt; http://www.usafa.af.mil/jscope/jscope99.html

[NAV98] Navin, Victor. 1998. *OLAP - A Thorough Overview*. Grattan Street, Parkville 3052, Australia: Department of Computer Science, The University of Melbourne, 30-Oct 1998. Seminar handout.

[NEB97] Nebesh, Bohdan Alexander. 1997. Using the World Wide Web for Experimentation in Reusable Component Comprehension. D.Sc Thesis, The George Washington University.

[NIE94] Nielsen, Jakob. 1994. *Usability Engineering*. Boston: Academic Press Professional.

[NIM98] NIMA. 1998. *Video Imagery Standards Profile (VISP) Version 1.3*. Washington, DC: National Imagery and Mapping Agency - Video Working Group (VWG), 6-Mar 1998. Report, VISP V3. http://164.214.2.51/vwg/std_prof/vidsprof.html

[NIS96] NIST. 1996. *Application Portability Profile (APP) The U.S. Government's Open System Environment Profile Version 3.0*.:, February 1996. Report, NIST SPEC PUB 500-230. http://www.nist.gov/itl/lab/list88/list88.htm

[NOR88] Norman, D. A. 1988. *The Psychology of Everyday Things*. New York: Basic Books.

[OSJ98] OS-JTF. 1998. *Open Systems - Joint Task Force Definitions*. Alexandria, VA: Office of the Under Secretary of Defense (Acquisition and Technology), 1-Apr 1998. Report. http://www.acq.osd.mil/osjtf/terms.htm

[OUE98] Ouellette, Tim. 1998. *Packaged middleware makes inroads: Off-the-shelf vs. custom-built*. Computerworld, April 13 1998. Web Site. http://www.computerworld.com/home/features.nsf/All/980413qs

[PAT90] Patton, Michael Quinn. 1990. *Qualitative Evaluation and Research Methods*. London: Sage Publications. http://union1.tui.edu/Faculty/FacultyGrad/Patton.htm

[PER94b] Perlman, Gary. 1994. The HCI bibliography: past, present, and future. In *ACM Conference on Human Factors in Computing Systems Celebrating Interdependence*: 71-72. Boston.

[PER98] Perlman, Gary. 1998. *Web-Based User Interface Evaluation with Questionnaires*. October 10 1998. http://www.acm.org/~perlman/question.html

[PIR74] Pirsig, Robert M. 1974. *Zen and the Art of Motorcycle Maintenance*. Quoted in The Columbia Dictionary of Quotations: Columbia University Press.

[PLA95] Plaice, John. 1995. Computer science is an experimental science. *ACM Computing Surveys* 27, no. 1: 33. http://www.acm.org/pubs/citations/journals/surveys/1995-27-1/p33-plaice/

[POR93] Porteous, M., J. Kirakowski, and M Corbett. 1993. *SUMI User Handbook*. Cork: Human Factors Research Group, University College Cork1993. http://143.169.20.1/MAN/WP51/t38.html

[RAG99] Raggett, Dave, Arnaud Le Hors, and Ian Jacobs. 1999. *HTML 4.01 Specification*. W3C, 24 August 1999. Web Site. http://www.w3.org/TR/html40/

[RED95] Reddy, Saveen. 1995. Should computer scientists worry about ethics? Don Gotterbarn says, ``Yes!''. *Crossroads* 1, no. 4. http://www.acm.org/crossroads/xrds5-4/dumbagents.html

[REN91] Rengger, Ralph. 1991. Indicators of Usability Based on Performance. In *Proceedings of the Fourth International Conference on Human-Computer Interaction*, ed. Bullinger, 1: 656--660.

[REY91] Reynolds, Jerry. 1991. If God had an agent, the world wouldn't be built yet. It'd only be about Thursday. *Quoted in: Newsweek. Quoted in The Columbia Dictionary of Quotations*, 25 November.

[RMO99] RM-ODP. 1999. *The Reference Model for Open Distributed Processing (RM-ODP)*. Distributed Systems Technology Centre, 1999. 1999. http://www-cs.open.ac.uk/ m_newton/odissey/RMODP.html; http://www-cs.open.ac.uk/ m_newton/odissey/RMODP.html; http://archive.dstc.edu.au/AU/research_news/odp/ref_model/ref_model.html

[ROS97] Rosenthal, Aaron, Eric Hughes, Scott Renner, and Len Seligman. 1997. A consumer viewpoint on `mediator languages -- a proposal for a standard`. *SIGMOD Record* 26, no. 1: 45-46. http://www.acm.org/sigmod/record/issues/9703/index.html

[ROU99a] Routio, Pentti. 1999. *Arteology: Ethical Considerations.* University of Art and Design Helsinki UIAH, March 8 1999. Web Site. http://www.uiah.fi/projects/metodi/151.htm

[RUB94] Rubin, Jeffrey. 1994. *Handbook of Usability Testing : How to Plan, Design, and Conduct Effective Tests (Wiley Technical Communication Library)*. New York: John Wiley and Sons.

[RUS95] Russell, Stuart and Peter Norvig. 1995. *Artificial Intelligence, A Modern Approach*. Upper Saddle River, NJ: Prentice-Hall.

[SAM97] Samadi, Shahin. 1997. Component Identification Framework and Toolkit (CIFT) In Support of Reusable Class Libraries. D.Sc Thesis, The George Washington University.

[SCH98] Schafer, Joseph H., Timothy J. Rogers, and John A. Marin. 1998. Networked Visualization of Heterogeneous US Army War Reserve Readiness Data. In *4th international Workshop on Advances in Multimedia Information Systems, MIS98*, ed. Sushil Jajodia, M. Tamer Ozsu, and Asuman Dogac, 1508. Istanbul, Turkey: Springer-Verlag.

[SCI89] Task Force on the Core of Computer Science. 1989. Computing as a Discipline. *Communications of the ACM* 32, no. 1: 23-Sep.

[SHA65] Shacklock, Richard. 1565. If many Faultes from "Epistle of Orosius" (Antwerp, 1565). In *Apple DOS Version 3.2 Manual*. Cupertino: Apple Computer Inc, 1979.

[SHN83] Shneiderman, Ben. 1983. Direct Manipulation: A Step Beyond Programming Languages. *IEEE Computer* 16, no. 8: 57-69.

[SHN93] Shneiderman, Ben. 1993. *Dynamic queries: for visual information seeking*. College Park: University of Maryland, January 1993. Report, CS-TR-3022. http://www.cs.umd.edu/projects/hcil/pubs/tech-reports.html; ftp://ftp.cs.umd.edu/pub/papers/papers/3022/3022.ps.Z

[SHN97a] Shneiderman, Ben. 1997. *Designing the User Interface : Strategies for Effective Human-Computer Interaction*. Reading, MA: Addison Wesley Publishing Company.

[SHN97b] Shneiderman, Ben. 1997. Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces. In *International Conference on Intelligent User Interfaces*: 33-39. http://www.acm.org/pubs/articles/proceedings/uist/238218/p33-shneiderman/p33-shneiderman.pdf

[SHN97c] Shneiderman, Ben and Pattie Maes. 1997. Direct Manipulation vs Interface Agents: Excerpts from debates at IUI 97 and CHI 97. *ACM interactions* IV, no. 6: 42-61. http://www.acm.org/pubs/articles/journals/interactions/1997-4-6/p42-shneiderman/p42-shneiderman.pdf

[SHO93] Shoham, Yoav. 1993. Agent-oriented programming. *Artificial Intelligence* 60: 51-92. http://www-csli.stanford.edu/csli/9495reps/interface9495-shoham.html

[SMI98a] Smith, Todd L. and Joseph H. Schafer. 1998. Using Machine Learning to Determine United States Army Readiness. In *Tenth Annual Software Technology Conference*. Salt Lake City, UT.

[SOL00] Solis, Gary. March 14 2000. Legal and Ethical Consequences of Research. Personal Email.

[SPS99] SPSS. 1999. *SPSS Base 10.0 User's Guide*. Chicago: SPSS Inc. www.spss.com

[SPY92] Spyridakis, Jan H. 1992. Conducting Research in Technical Communication: The Application of True Experimental Designs. 39, no. 4: 607-624. http://www.stc-va.org/fstcperiodicals.htm

[STE94] Sternstein, Martin. 1994. *Statistics*. Hauppauge, NY: Barron's Educational Series, Inc.

[STO97b] Stoffel, Kilian, Merwyn Taylor, and James A. Hendler. 1997. Efficient Management of Very Large Ontologies. In *American Association for Artificial Intelligence Conference*: AAAI / MIT Press.

[SUB00] Subrahmanian, V. S., Thomas Eiter, Piero Bonatti, Sarit Kraus, Robert Ross, Jürgen Dix, and Fatma Ozcan. 2000. *Heterogeneous Agent Systems*: MIT Press. http://mitpress.mit.edu/book-home.tcl?isbn=0262194368

[SUB97a] Subrahmanian, V. S., Sibel Adali, Anne Brink, James J. Lu, Adil Rajput, Timothy J. Rogers, Robert Ross, and Charles Ward. 1997. *HERMES: A Heterogeneous Reasoning and Mediator System*. College Park: University of Maryland1997. Report. http://www.cs.umd.edu/projects/hermes/overview/overview.html; http://www.cs.umd.edu/projects/hermes/overview/paper/index.html

[SUB99a] Subrahmanian, V. S. 1999. Heterogeneous agent systems. In *Proceedings of the 11th International Symposium on Foundations of Intelligent Systems (ISMIS99)*, ed. Zbigniew W. Ra and Andrzej Skowron, 1609: 46--55. Berlin: Springer.

[SUB99b] Subrahmanian, V. S. 1999. Heterogeneous agent systems. *Lecture Notes in Computer Science* 1609: 46--55.

[SUT91] Sutcliffe, Alistair, John Carroll, Richard Young, and John Long. 1991. HCI theory on trial. In *ACM Conference on Human Factors in Computing Systems Reaching Through Technology*: 399-401. http://www.acm.org/pubs/citations/proceedings/chi/108844/p399-sutcliffe/

[SUT98] Sutter, James R. 1998. Project-based warehouses. *Communications of the ACM* 41, no. 9: 49-51. http://www.acm.org/pubs/citations/journals/cacm/1998-41-9/p49-sutter/

[SWE98] Sweeney-Jackson, Lethia. 1998. Developing and Validating an Instrument to Analyze the Legibility of a Web Page Based Upon Text Style and Color Combinations. D.Sc Thesis, The George Washington University.

[SYC98a] Sycara, Katia P. 1998. The Many Faces of Agents. *AI Magazine* 19, no. 2: 12-Nov.

[TAN98] Tannen, Val. 1998. Heterogeneous Data Integration with Mobile Information Manager. *Lecture Notes in Computer Science* 1508: 2.

[TAY97] Taylor, Merwyn, Kilian Stoffel, and James A. Hendler. 1997. Ontology-based Induction of High Level Classification Rules. In *SIGMOD Data Mining and Knowledge Discovery Workshop*. Tucson, AZ.

[THO97] Thompson, Jess. 1997. Toolbox: Avoiding a Middleware Muddle. *IEEE Software* 14, no. 6: 92-95. http://dlib.computer.org/so/books/so1997/pdf/s6092.pdf

[TIC98] Tichy, Walter F. 1998. Should Computer Scientists Experiment More? *IEEE Computer* 31, no. 5: 32-40.

[TRO99] Trochim, William M. 1999. *The Research Methods Knowledge Base, 2nd Edition*. 9-Apr 1999. Web Site. http://trochim.human.cornell.edu/kb

[TSP98] TSP158-A-3000. 1992. *Training Support Package 158-A-3000, Apply the Brigade Fight Concept to Operational Planning*. Washington, DC: Headquarters, Department of the Army. http://www.adtdl.army.mil/cgi-bin/atdl.dll/cctsp/158-a-3000/158-a-3000.htm

[TUF90] Tufte, Edward R. 1990. *Envisioning Information*. Cheshire, CT: Graphics Press.

[TUL85] Tullis, Thomas S. 1985. Designing a Menu-Based Interface to an Operating System. In *Proceedings of ACM CHI'85 Conference on Human Factors in Computing Systems*: 79--84.

[TUL93] Tullis, Thomas S. 1993. Is User Interface Design Just Common Sense? In *Proceedings of the Fifth International Conference on Human-Computer Interaction*, 2: 9--14.

[TUL95] Tullis, Thomas S. and Marianne L. Kodimer. 1995. A Comparison of Direct-Manipulation, Selection, and Data-Entry Techniques for Reordering Fields in a Table. In *Human Factors Perspectives on Human-Computer Interaction: Selections from Proceedings of Human Factors and Ergonomics Society Annual Meetings, 1983-1994*, ed. G. Perlman, G. K. Green, and M. S. Wogalter: 266-270. Santa Monica, CA: Human Factors and Ergonomics Society.

[VAN97a] van Dam, Andries. 1997. The Human Connection: Post-WIMP User Interfaces. *Communications of the ACM* 40, no. 2: 63-67. http://www.acm.org/pubs/citations/journals/cacm/1997-40-2/p63-van_dam

[VEE95] Veerasamy, Aravindan. 1995. Querying, Navigating and Visualizing an Online Library Catalog. In *18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*: 367. http://www.csdl.tamu.edu/DL95/papers/veerasamy/veerasamy.html

[VIR92] Virzi, Robert A. 1992. Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? *Human Factors* 34, no. 4: 457--468. http://www.hcibib.org/gs.cgi?word=checked&terms=C.HFS.90.291

[WAL98] Waltz, Edward. 1998. *Information Warfare: principles and operations*. Boston: Artech House.

[WAT97] Watson, Mark. 1997. *Intelligent Java Applications for the Internet and Intranets*. San Francisco: Morgan Kaufmann Publishers. http://www.amazon.com/exec/obidos/ISBN=1558604200

[WEG96] Wegner, Peter and Jon Doyle. 1996. Editorial: strategic directions in computing research. *ACM Computing Surveys* 28, no. 4: 565-574. http://www.acm.org/pubs/citations/journals/surveys/1996-28-4/p565-wegner/

[WEL37] Wells, Carolyn. 1937. *The Rest of My Life*. Quoted in The Columbia Dictionary of Quotations: Columbia University Press.

[WHI98] White, Bobby. 1998. Logistics Integration Agency Army War Reserve System. Workshop at New Cumberland Defense Distribution Center.

[WIE92] Wiederhold, Giovanni Corrado Melchiore. 1992. Mediators in the Architecture of Future Information Systems. *IEEE Computer* 25, no. 3: 38-49.

[WOO95b] Wooldridge, Michael J. and N. R. Jennings. 1995. Intelligent Agents: Theory and Practice. *Knowledge Engineering Review* 10, no. 2. http://www.elec.qmw.ac.uk/dai/people/mikew/pubs/ker95/ker95-html.html

[YEE91] Yee, M. 1991. System design and cataloging meet the user: User interfaces to online public access catalogs. *Journal of the American Society for Information Science* 42, no. 2: 78-98.

[YOU94] Young, J and K. Maracaccio. 1994. *Gale Directory of Databases*. Detroit: Gale Research, Inc.

[ZIE96] Ziegler, Jürgen. 1996. Interactive techniques. *ACM Computing Surveys* 28, no. 1: 185-187. http://www.acm.org/pubs/citations/journals/surveys/1996-28-1/p185-ziegler/

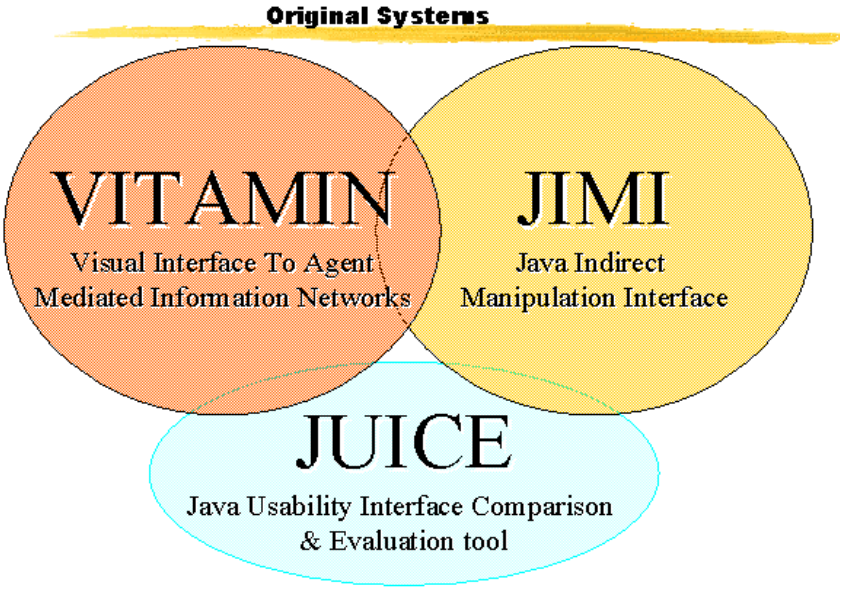
APPENDIX A: SUBJECT INSTRUCTIONS

This appendix contains a copy of the research web site used for the experiments. It has been reformatted to fit this paper.

Hello and welcome!

Background

The purpose of this experiment is to evaluate and compare two different Java interfaces to Army War Reserve data. The interfaces allow Army planners and leaders to understand the current state of readiness of prepositioned equipment aboard ships. The data required to present this information is in many old (legacy) and scattered (distributed) databases. An additional purpose of this experiment is to gather your input on the program that enables this experiment.



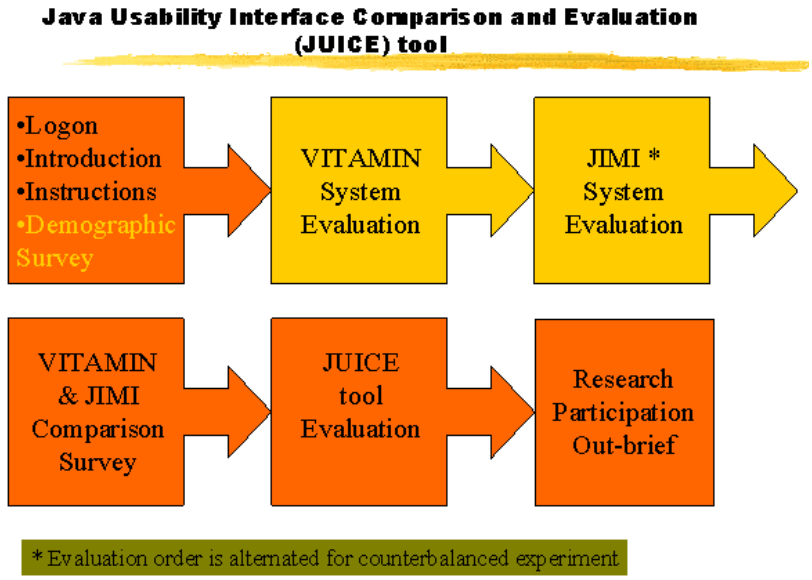
The two interfaces are called JIMI and VITAMIN. The program or tool that will be used to conduct the experiment is called JUICE.

JIMI is the Java Indirect Manipulation Interface. VITAMIN is the Visual Interface to Agent Mediated Information Networks. JUICE is the Java Usability Interface Comparison and Evaluation tool.	

SEQUENCE

After this introduction and completion of the "Informed Consent Before Participation" sheet, you will begin the experiment with the logon information provided by the instructor.

Then the JUICE tool will take you through the following:

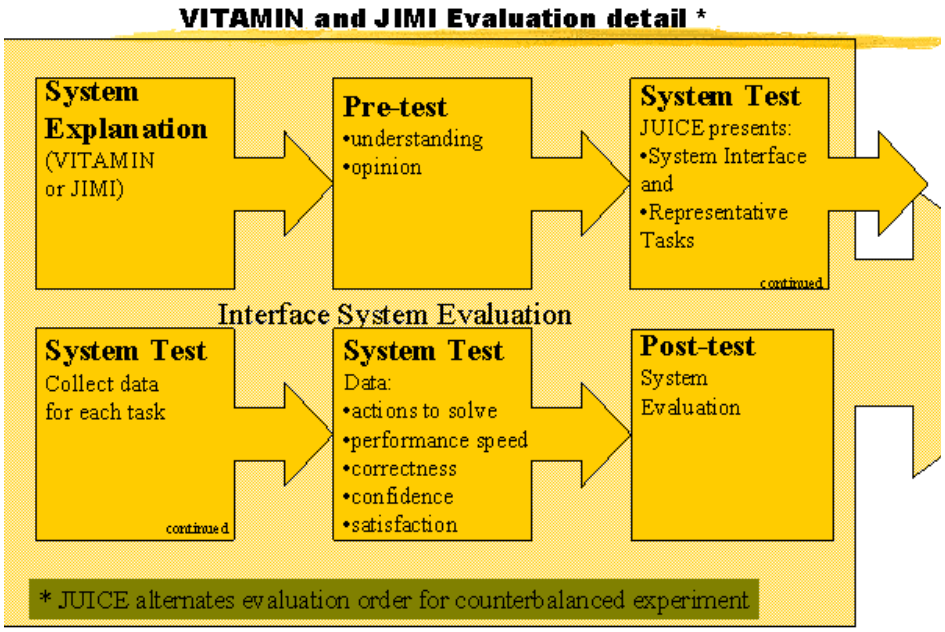


• logon, brief instructions, and a demographic survey • an evaluation of JIMI or VITAMIN based on representative tasks (order determined randomly by JUICE) • an evaluation of the other interface (JIMI or VITAMIN) based on representative tasks • a comparison survey of VITAMIN & JUICE • a formative evaluation of the JUICE tool itself • logout and thank you	

Then you will complete the "Informed Consent After Participation" sheet and receive a "Research Participation Worksheet" for your PL100 instructor.

The entire experiment will take 55 minutes. Work quickly. Once you have answered a question and hit the "Next" button, there is no way to return and change your answer.

Interface Evaluation Detail



JUICE will take you through an evaluation of VITAMIN and JIMI	

JUICE will display the interface (JIMI or VITAMIN) on the left and JUICE will display its instructions and questions on the left.

Below are non-active images from the interfaces with detailed instructions and background. Please study then carefully and quickly before you begin.

VITAMIN INSTRUCTIONS

When VITAMIN first comes up the data and query window will be blank.
Hit the "Populate" button on the bottom to send off the first query this query will be a loc_Totals query and will return the list of ships (SzLOC) and three columns with the Authorized_Quantity, OnHand_Quantity, and Percent_Fill. (You should only have to press the Populate button once during your evaluation of VITAMIN.)
For the afloat sets the location is the same as the ship name.
When you hit the Plus buttons then two more queries will be generated for that ship - loc_ERC_Totals and loc_force_totals. The results will be displayed under the ship and you will "drill into" the data.

"NEW CUMBERLAND"	7697	6971	90.5678
"ALEXANDRIA"	9510	8471	89.0747
"WEST POINT"	8402	7701	91.6567
"COLLEGE PARK"	8057	7170	88.9909
"ARLINGTON"	0	0	0
"CAMP HILL"	0	0	0
"ARKANSAS"	0	0	0
"KEY WEST"	384	342	89.0625
"SLEEPY HOLLOW"	1219	1099	90.1559

SzLOC, Authorized_Quantity, OnHand_Quantity, Percent_Fill	ERC is the Equipment Readiness Code - "P" = pacing items for instance in a Tank Company, the Tanks would be the Pacing Items. Other ERCs are "A" and "B/C" Each Ship has a Brigade Size set of equipment with three force types: "BN TF" is the combat Battalion Task Force "BDE CS/CSS" is the Brigade Combat Support and Combat Service Support "EAD CS / CSS" is the Echelons Above Division CS/CSS Both the ERC and Force queries will have Plusses to get information down to the Company level - this level is called "DUIC" - for Derivative Unit Identification Code. The queries (or "Predicates") are loc_ERC_DUIC_totals, and loc_force_DUIC_totals, respectively. Each level will list the Authorized, OnHand, and % fill and the DUIC level will also have a more detailed description.

You will answer the task questions by submitting queries by hitting the plus buttons. Hitting the minus button will collapse the query.
The button color is determined by the D_Status based on %fill. Above 90% D_Status = 3 and we see a green button, above 80% D_Status is 2 and the button is Yellow, below 80% D_Status is 1 and the button is red.

JIMI INSTRUCTIONS

With JIMI, your first query should also be the loc_Totals so you can get an overview of the data. To do the query with JIMI, double click on the loc_Totals predicate, this will place a copy of the query text into the Query text area. Then just hit the Execute Query button. (See the box below for the result.)

All of the predicates for JIMI are the same as those for VITAMIN. The data sets are different - so you will notice different ship (SzLOC) names.	A feature of JIMI is the ability to edit the query text before you execute the query. For instance, you can replace the variable SzLOC with the name of one of the ships i.e. "OSHKOSH". This will let you query about just that one ship. Remember that the names are in ALL CAPS and surrounded by "" (double quotes). Variables you may wish to fill in are SzLOC, SzERC, and SzForce. See below for explanations and legal values. **IMPORTANT POINTS:** *- Some queries require one or more variables to be specified - this will be indicated in the Description box.* *- Do not submit one of the DUIC queries without filling in one of the variables. (SzLOC is required to be filled in).* These actions will result in a very long wait while the mediator agent attempts to resolve the query.

The image below shows the results of the loc_Totals predicate search (or query) with none of the variables specified. You can see that nine answers were returned. To see the answer text, click on the number of the answer you wish to investigate. Other variables you will see are: SzERC - is a sting for the ERCs: "P", "A", or "B/C" SzForce (or SzFrc) - is a string for the subordinate Force units on the ship. Legal values are: "BN TF", "BDE CS/CSS" and "EAD CS/CSS" The DUIC queries also have a SzDesc, which returns a description string for the company.	Notice that the answer format is the same as for VITAMIN - and you can verify it by looking at the Query Text. • D_Status represents the readiness rating (using a simple unclassified algorithm). (D means that is a "Double" precision floating point number). • SzLOC is the Location or Ship name (Sz means "String" of characters). • D_AuthQty is the Authorized Quantity of Equipment • D_OnHand is the OnHand Quantity of Equipment • D_Percent is the % fill (Auth/OnHand) of Equipment

Please read all of the above carefully and then BEGIN

APPENDIX B: SUBJECT BACKGROUND SURVEY

B.1 Dissertation experiment Survey

1. How long have you been using computers?
A. less than a year
B. one to two years
C. two to five years
D. more than five years

2. Choose your highest education level?
A. High School Graduate
B. College Student
C. College Graduate
D. Graduate Degree

B.2 Pilot Study Survey

1. How long have you been using computers?
A. less than a year
B. one to two years
C. two to five years
D. more than five years
2. What introduction to Computer Science course have you taken?
A. CS105 Spring 1999
B. CS155 Spring 1999
C. CS105 Fall 1998
D. CS155 Fall 1998
3. How many different types of computer systems have you used (e.g., MAC, IBM-PC DOS, Sun, Windows)?
A. one
B. less than four but one or more
C. less than nine but four or more
D. more than nine
4. How many different computer packages you probably used (e.g., Excel, Word)?
A. only one
B. less than five but more than one
C. less than ten but more than five
D. more than ten
5. How long do you use a computer each week (total hours)?
A. only one hour or less (10 - 60 minutes)
B. about a days work for each week (1 - 8 hours)
C. more than a days work (8 - 30 hours)
D. all week, every day (more than 30 hours)

6. How long do you use a computer each week (on the world wide web)?
A. only one hour or less (10 - 60 minutes)
B. about a days work for each week (1 - 8 hours)
C. more than a days work (8 - 30 hours)
D. all week, every day (more than 30 hours)
7. Do you regularly use any web search engines (e.g., yahoo, Alta vista) or any other database search tools?
A. not at all (only 1 or 2 times perhaps)
B. from time to time (not everyday but more than 2 in my life)
C. regularly enough (once everyday)
D. frequently (more than once everyday)
8. Do you regularly use any database search tools?
A. not at all (only 1 or 2 times perhaps)
B. from time to time (not everyday but more than 2 in my life)
C. regularly enough (once everyday)
D. frequently (more than once everyday)
9. Do you use any data analysis tools (e.g., Excel) ?
A. not at all (only up to 1 or 2 times in my life)
B. from time to time (not everyday but more than 2 in my life)
C. regularly (once everyday)
D. frequently (more than once everyday)

APPENDIX C: EXPERIMENTAL TASK LISTS

All experimental task questions are five item multiple-choice. There is one set of questions for VITAMIN and one for JIMI. The only difference between the two sets of questions is the ship names.

C.1 Confidence question for each task

How confident are you in this answer?	Analysis Value
A. Extremely	1.00
B. Very	0.75
C. Neutral	0.50
D. Not Very	0.25
E. Not At All	0.00

Table 33 Confidence Values

C.2 VITAMIN Task Questions

Example 1. How many ships are there overall?
A. 2
B. 3
C. 4
D. 5
E. 6 //correct answer

Example 2. What is the Percent Fill of Pacing Items (SzERC= P) on the ship named ALEXANDRIA?
A. 95.045 //correct answer
B. 79.792
C. 444
D. 89.0747
E. 0

Example 3. How many total Infantry and Tank Companies are in the BN TF on the ship named ALEXANDRIA?
A. 2
B. 3
C. 4 //correct answer
D. 5
E. 0

4. Which SHIP has the highest overall Percent Fill (D_Percent)?
A. SLEEPY HOLLOW
B. WEST POINT //correct answer
C. COLLEGE PARK
D. KEY WEST
E. NEW CUMBERLAND

5. What is the Percent Fill for all items on the ship named SLEEPY HOLLOW?
A. 90.1559 //correct answer
B. 89.9909
C. 90.5678
D. 89.0625
E. 0

6. Which ship has the smallest overall authorized quantity of equipment?
A. SLEEPY HOLLOW
B. WEST POINT
C. COLLEGE PARK
D. KEY WEST //correct answer
E. NEW CUMBERLAND

7. What is the overall Authorized Quantity for all items in the SzFrc = BDE CS/CSS slice on the ship named NEW CUMBERLAND?
A. 4705
B. 9826
C. 2018 //correct answer
D. 5456
E. 0

8. What is the Percent Fill of the SzFrc = BN TF on the ship named COLLEGE PARK?
A. 92.208
B. 23.53
C. 89.5358 //correct answer
D. 98.0154
E. 0

9. How many SzERC = A items are ON HAND on the SLEEPY HOLLOW?
A. 163 //correct answer
B. 207
C. 2
D. 1010
E. 934

10. Which ship has the highest Percent Fill for SzERC A items?
A. SLEEPY HOLLOW
B. NEW CUMBERLAND //correct answer
C. COLLEGE PARK
D. KEY WEST
E. ALEXANDRIA

11. How many Tank Companies are in the BN TF on the ship named WEST POINT?
A. 0
B. 1
C. 2 //correct answer
D. 3
E. 4

12. How many Medical Companies are in the BDE CS/CSS on the ship named COLLEGE PARK?
A. 0
B. 1 //correct answer
C. 2
D. 3
E. 4

C.3 JIMI Task Questions

Example 1. How many ships are there overall?
A. 2
B. 3
C. 4
D. 5
E. 6 //correct answer

Example 2. What is the Percent Fill of Pacing Items (SzERC= P) on the ship named WHEELING?
A. 95.045 //correct answer
B. 79.792
C. 444
D. 89.0747
E. 0

Example 3. How many total Infantry and Tank Companies are in the BN TF on the ship named WHEELING?

A. 2
B. 3
C. 4 //correct answer
D. 5
E. 0

4. Which SHIP has the highest overall Percent Fill (D_Percent)?
A. CAMP BUCKNER
B. BROADVIEW HTS //correct answer
C. FOGGY BOTTOM
D. HIGHLAND FALLS
E. OSHKOSH

5. What is the Percent Fill for all items on the ship named CAMP BUCKNER?
A. 90.1559 //correct answer
B. 89.9909
C. 90.5678
D. 89.0625
E. 0

6. Which ship has the smallest overall authorized quantity of equipment?
A. CAMP BUCKNER
B. BROADVIEW HTS
C. FOGGY BOTTOM
D. HIGHLAND FALLS //correct answer
E. OSHKOSH

7. What is the overall Authorized Quantity for all items in the SzFrc = BDE CS/CSS slice on the ship named OSHKOSH ?
A. 4705
B. 9826
C. 2018 //correct answer
D. 5456
E. 0

8. What is the Percent Fill of the SzFrc = BN TF on the ship named FOGGY BOTTOM?
A. 92.208
B. 23.53
C. 89.5358 //correct answer
D. 98.0154
E. 0

9. How many SzERC = A items are ON HAND on the CAMP BUCKNER ?
A. 163 //correct answer
B. 207
C. 2
D. 1010
E. 934

10. Which ship has the highest Percent Fill for SzERC A items?
A. CAMP BUCKNER
B. OSHKOSH //correct answer
C. FOGGY BOTTOM
D. HIGHLAND FALLS
E. WHEELING

11. How many Tank Companies are in the BN TF on the ship named BROADVIEW HTS?
A. 0
B. 1
C. 2 //correct answer
D. 3
E. 4

12. How many Medical Companies are in the BDE CS/CSS on the ship named FOGGY BOTTOM?
A. 0
B. 1 //correct answer
C. 2
D. 3
E. 4

C.4 Ship Name Mappings

JIMI	**VITAMIN**
Washington	Arkansas
Broadview Heights	West Point
Foggy Bottom	College Park
Highland Falls	Key West
Wheeling	Alexandria
New Cumberland	Oshkosh

Table 34 Ship Name Mappings

APPENDIX D: USER PREFERENCE QUESTIONNAIRES

D.1 Preference Categories

Easy to use
Helpful
Perform Faster
Information Quality
Interface Quality
Learn more about the data
Enjoyable
Useful
Comments

D.2 Survey Comparison Questions

Select one response to each question.
A. Strongly Disagree
B. Disagree
C. Neither Agree no Disagree
D. Agree
E. Strongly Agree

1. I found the VITAMIN System GUI to the agent-mediated legacy data easy to use

2. I found the Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data easy to use.

3. I found the VITAMIN System GUI to the agent-mediated legacy data helpful

4. I found the Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data helpful

5. The VITAMIN System GUI to the agent-mediated legacy data allowed me to perform faster

6. The Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data allowed me to perform faster

7. The VITAMIN System GUI to the agent-mediated legacy data provided high information quality

8. The Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data provided high information quality

9. The VITAMIN System GUI to the agent-mediated legacy data provided high interface quality

10. The Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data provided high interface quality

11. The VITAMIN System GUI to the agent-mediated legacy data allowed me to learn more about the data

12. The Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data allowed me to learn more about the data

13. I found the VITAMIN System GUI to the agent-mediated legacy data enjoyable to use

14. I found the Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data enjoyable to use

15. I found the VITAMIN System GUI to the agent-mediated legacy data useful.

16. I found the Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data useful

17. Comments about the VITAMIN System GUI.

18. Comments about the Java Indirect Manipulation Interface (JIMI).

19. Overall Comments

D.3 Survey Analysis Scale

Multiple Choice Selection	Analysis Value	Sum Analysis Value
A. Strongly Disagree	0.00	0.0
B. Disagree	0.25	2.0
C. Neither Agree no Disagree	0.50	4.0
D. Agree	0.75	6.0
E. Strongly Agree	1.00	8.0

Table 35 Survey Analysis Scale

Appendix E: JUICE SYSTEM EVALUATION SURVEY QUESTIONS

This survey was used for the pilot studies. It was not used for the dissertation experiment reported in this thesis.

1. The JUICE tool succeeded in enabling my evaluation of the two interfaces.
A. Strongly Disagree
B. Disagree
C. Neither Agree no Disagree
D. Agree
E. Strongly Agree
2. The Juice tool succeeded in attracting and maintaining my interest during the evaluation period.
A. Strongly Disagree
B. Disagree
C. Neither Agree no Disagree
D. Agree
E. Strongly Agree
3. I found the JUICE tool to be easy to use.
A. Strongly Disagree
B. Disagree
C. Neither Agree no Disagree
D. Agree
E. Strongly Agree
4. I found the JUICE tool to present a practical method for comparing two interfaces.
A. Strongly Disagree
B. Disagree
C. Neither Agree no Disagree
D. Agree
E. Strongly Agree
5. I believe that it is important to evaluate and compare user interfaces.
A. Strongly Disagree
B. Disagree
C. Neither Agree no Disagree
D. Agree
E. Strongly Agree
6. The Java Indirect Manipulation Interface (JIMI) GUI to the agent-mediated legacy data allowed me to perform faster
A. Strongly Disagree B. Disagree C. Neither Agree no Disagree D. Agree E. Strongly Agree
7. JUICE Comments - What Works?
8. JUICE Comments - What needs Work?
9. JUICE Comments - How can it be improved?

APPENDIX F: HUMAN RESEARCH POLICIES

This Appendix contains the procedures, explanations, and forms for human research at USMA.

F.1 Procedure to Use Cadets as Participants in Research

1. It will take about two weeks after submission of necessary forms to the Research Coordinator to receive approval from the Human Subjects Committee. All the necessary forms are in the **G:\\z-ranger\public\subjects** directory. These are read-only/copy-only files. These are also available at http://www-internal.dean.usma.edu/bsl/research.htm. Please note that the name and signature sections for the principal researcher/supervisor in the forms are for faculty members/researchers in the department, not for cadets.

The principal researcher / supervisor (Note: Hereon this will be referred to as the researcher.) will fill out the SUBJECT USE REQUEST FORM and submit it to the Research Coordinator - Dr. Sehchang Hah, (914) 938-5628, Room TH262. All research must be carried out in the BS&L area.

2. The Research Coordinator will check the content of the submitted form to insure it adheres to the guidelines established by American Psychological Association (APA) and USMA.
3. If it needs modifications, it will be returned to the researcher.
4. If it is accepted, the researcher will be notified by the Research Coordinator to submit a Memorandum to the Chairman of the Human Subjects Committee through the Research Coordinator, the Director of Psychology Programs, and the Department Head for the final approval. A standard memorandum form is available in the above-mentioned directory.
5. The researcher whose proposal is accepted will get participant-hours allocated by the Research Coordinator.
6. The researcher will submit the RESEARCH SIGN-UP SHEET to the Research Coordinator to get the research number. The sign-up sheet will have 30 minutes or longer time-slots for cadets to sign up and descriptions of cadets' tasks to perform. This form will be posted on the wall in the hallway near Room 259, Thayer Hall. The researcher may ask a cadet to contact him/her by E-mail and the researcher and the cadet may schedule electronically. <u>In this case, the researcher must put the cadet's name in the proper time-slot in the RESEARCH SIGN-UP SHEET to keep the record.</u>
7. The researcher will get signatures from a cadet both before and after the experiment by having him/her complete the INFORMED CONSENT form. The researcher will make a copy of the INFORMED CONSENT form and give it to the cadet for his/her information.
8. The researcher will discuss the questions in the WORKSHEET with his/her cadet during debriefing. The cadet will fill out the form at his/her convenience and submit it to his/her instructor as soon as possible. The instructor will keep track of his/her cadets' participation using the information in the WORKSHEET.
9. If the signed-up cadet does not show up, the researcher will inform of the fact to his/her instructor. The instructor will ask the cadet to sign up again. If the cadet shows up but the researcher does not, the researcher fills out a worksheet and sends it to the cadet electronically. The researcher will notify the cadet's instructor of the incident. The cadet will turn in the worksheet to the instructor to get credit.
10. The researcher will submit both the RESEARCH SIGN-UP SHEET and the INFORMED CONSENT forms to the Research Coordinator by the Lesson 38 of the semester during which the research is performed.

F.2 SUBJECT USE REQUEST FORM

Semester/Year: Spring 1999
Research Number: USMA BS&L Spring 1999 Number 5
(Office Use Only)

1. Title of the research: Evaluating Two Visual Interfaces to Mediated War Reserve Readiness Data
2. Experimenter(s)/Data Collector(s): MAJ Schafer, (914) 938-2407, joseph-schafer@usma.edu (dj4149)
3. Principle Researcher(s)/Supervisor(s): MAJ Joseph H. Schafer
4. Summary of the research:
Thirty subjects will each spend forty-five minutes on a controlled psychologically oriented usability experiment. The experiment will use a newly developed automated remote interface usability-testing tool to compare two interfaces in a within-subject, counterbalanced procedure. Dependent measures include performance times on representative tasks, error rates, subjective user satisfaction, and confidence. Subjects will also offer an appraisal of the usability-testing tool.

5. Subject hours needed: Thirty
6. Anticipated starting date when the research will be conducted:
7. Place where the research will be conducted: Thayer Hall, West Point, NY
8. Telephone number for cancellations: (914) 938-2407 – MAJ Schafer
9. Purpose of request (Mark one):
   a. Class Project (XX): What Class? (CS408)
   b. Class Exercise ( ): What Class? ( )
   c. General Research ( ) d. Grant Research ( )
   e. Joint Research (with whom and what agency?: Include Name, Address, and Phone number) ( )
   f. Other (Please specify.) _____

The design and implementation of the usability-testing tool is part of a capstone design project in CS408. This tool, in turn, is part of a larger project to prototype and evaluate the Visual Interface To Agent Mediated Information Networks (VITAMIN) system as applied to distributed Army war reserve legacy data sources for the Logistics Integration Agency.

10. Brief description of procedure (include sufficient details regarding potentially hazardous, uncomfortable, or stressful aspects of the procedure):

Log onto a web site and answer multiple-choice questions. We anticipate no hazards, discomfort, or stress.

11. Is physical or psychological stress a necessary part of the research (such as sense of insecurity or failure, assault on values, electric shock, etc.)?

Yes ____ No __XX__

12. If you answered No in the above Question 11, skip to Question 13. Otherwise, please describe, in detail, the stress and possible harm from participating in this research.

13. Before participation, the Experimenter/Data Collector will brief cadets about the tasks to be performed. During this briefing, cadets will be told that they may stop participating in the research at anytime without penalty. The cadet will receive a participation credit even if s/he withdraws before completing the designed participation. Is Experimenter/Data Collector aware of this policy?

Yes __XX__ No ____

14. Is Experimenter/Data Collector aware that he/she should not cancel the research sessions in which cadets are scheduled except in cases of emergency?

Yes __XX__ No ____

15. Is Experimenter/Data Collector aware that if the session lasts more than 55 minutes, it should be counted as 2 credits? (Note: 3 credits will be given if participation requires over 110 minutes. Participation may not exceed 165 minutes without special permission.)

Yes __XX__ No ____

16. The personal identity of cadets will not be revealed in any form in the collected data or in the research report. Cadets' privacy must be respected. Is Experimenter/Data Collector aware of this right to privacy?

Yes __XX__ No ____

17. Please describe what educational gain cadets will receive by participating in this research.

The Army's Advanced Warfighting Experiments revealed in part that highly digitized forces can gain a tremendous advantage in situational awareness over their adversaries. Unfortunately, though the digitization proved a great value before and after the fight, in the heat of battle, among other issues, soldiers and commanders did not want to trust and wait for a display interface to update. They fell back to manual procedures.
This research compares two visual interfaces to Army war reserve readiness data. Cadets will gain an appreciation for the complexity of using information from many different sources to answer simple queries and they will gain a deeper understanding of aspects of different interface designs.
Additionally, cadets will have an opportunity to interact with and provide feedback on an automated remote usability-testing tool similar to systems, which may have a significant impact on their digitized Army.

18. The sign-up sheets, with the stated procedure, are to be returned to the Research Coordinator. They will be stored in the department for two years as required by NIMH. Informed Consent Forms are also to be returned to the Research Coordinator. Is Experimenter/Data Collector aware of these policies?

Yes __XX__ No ____

19. Please attach all material to be presented to cadets, either orally or in writing, including Questionnaire, Debriefing Materials, and Instructions.

Subject Instructions, Subject Background Survey, Experimental Task Lists, User Preference Questionnaire draft materials enclosed.

20. My responsibility as Experimenter/Data Collector is clear to me.

Experimenter (s)/Data Collector(s)

Name (print) Joseph H. Schafer, MAJ USA (Signature)

21. I have discussed the proper procedure with the Experimenter(s)/ Data Collector(s) and will provide the close supervision to ensure safety and ethical standards are maintained.

Principle Researcher(s)/Supervisor(s):

(Print) Joseph H. Schafer, MAJ USA (Signature)

F.3 Application for Use of Human Participants in Research Memorandum

DEPARTMENT OF THE ARMY
UNITED STATES MILITARY ACADEMY
West Point, New York 10996

REPLY TO ATTENTION OF

MADN-AI 6 MARCH 1999

MEMORANDUM THRU

Dr Sehchang Hah, Research Coordinator, Department of Behavioral Sciences and Leadership

LTC Jose A. Picart, Academy Professor and Director, Psychology Programs, Department of Behavioral Sciences and Leadership

COL Charles F. Brower IV, Professor and Head, Department of Behavioral Sciences and Leadership

For COL Kerry Pierce, Director, Office of Policy, Planning, And Analysis

SUBJECT: Application for Use of Human Participants in Research

1. Title of the research: Evaluating Two Visual Interfaces to Mediated War Reserve Readiness Data

2. Principal researcher: MAJ Joseph Schafer

3. Objectives of research: Compare two interfaces and investigate a usability-testing tool.

4. Methods and procedure to be used (Include the approximate number of cadets needed and the estimated amount of time each cadet will volunteer for the research):

Thirty subjects will each spend forty-five minutes on a controlled psychologically-oriented usability experiment. The experiment will use a newly developed automated remote interface usability-testing tool to compare two interfaces in a within-subject, counterbalanced procedure. Dependent measures include performance times on representative tasks, error rates, subjective user satisfaction, and confidence. Subjects will also offer an appraisal of the usability-testing tool.

5. Research Number: USMA BS&L Spring 1999 Number 5

ORIGINAL SIGNED
JOSEPH H. SCHAFER
MAJ, SC

Office of Artificial Intelligence, Analysis, and Evaluation
Department of Electrical Engineering and Computer Science

F.4 PL100 Lesson 40 "Research Policy"

http://www-internal.dean.usma.edu/bsl/lsn40_policy.htm

1. This page outlines the policy concerning PL100 cadet participation in research conducted or sponsored by faculty in the Department of Behavioral Sciences and Leadership.
2. PL100 cadets can be excused from PL100 Lesson 40 and earn a maximum of 20 extra credit points by participating in departmental research. A cadet must actually participate in a research study to be excused from Lesson 40. For those cadets who attempt to participate but are unable to participate due to circumstances beyond the cadet's control (see paragraphs 4 and 5 below), the instructor will award a maximum of 20 extra credit points as described below. Cadets will NOT be excused from Lesson 40, however, unless they actually participate in a research study. PL100 extra credit points and excusal from Lesson 40 will be awarded as follows:
    a. By volunteering for, and participating in, one research study cadets will be excused from attending Lesson 40 (Research Laboratory Exercise). **PL100 Lesson 40 is a research laboratory exercise that will require cadets who have not participated in departmental research to read a psychology journal article on reserve in the USMA Library and submit a completed worksheet prior to Lesson 40. In addition, cadets who do not participate in departmental research will attend class on Lesson 40 and participate in a simulated research exercise.**
    b. By volunteering for a second research study, cadets will earn **eight (8) extra credit points** (the equivalent of 2 DSQs).
    c. By volunteering for a third research study, cadets will earn **twelve (12) additional extra credit points** (the equivalent of 3 more DSQs).
3. Research participation credit will be awarded based on the number of studies (lasting 55 minutes or less) in which the cadet participates. Cadets who participate in a study lasting more than 55 minutes, (but less that 110 minutes), will receive credit for participating in two studies (i.e., drop for lesson 40 and 8 extra credit points). In general, departmental research will not require more than 110 minutes of cadet participation in a single study.
4. If a cadet reports for a scheduled research study, but is unable to participate in the research study for any reason that is not the cadet's fault, the instructor will:
    a. verify that the cadet was present but was unable to participate through no fault of the cadet.
    b. award the cadet research participation extra credit points (either 8 or 12 points depending on the number of research studies the cadet has volunteered to participate in). The cadet will NOT receive a drop in lesson 40, however, unless he or she actually participates in a research study.
5. If a cadet reports to the correct place and at the correct time for a scheduled research study and the researcher is not present, cadets will take the following actions:
    a. Cadets will ask any BS&L faculty in the area for assistance in locating the researcher.
    b. If the cadet and the faculty member (or the cadet alone) are unable to locate the researcher within ten (10) minutes of the scheduled research start time, the cadet will sign-out on the sheet located at the research sign-up board. The cadet will enter the research study number, print their name, sign their name and enter the time they departed.
    c. Upon confirming that the cadet signed-out after making an effort to locate the absent researcher, the instructor will award the cadet research participation extra credit points as described in paragraph 4b. above. The cadet will NOT be excused from Lesson 40, unless he or she actually participates in a research study.
6. Cadet participation is VOLUNTARY. Although participation in research for extra credit will add points to the PL100 grade of those cadets who volunteer to participate, this will not effect the grades of those cadets who elect not to participate.

APPENDIX G: DISSERTATION EXPERIMENT FORMS

G.1 RESEARCH SIGN-UP SHEET[2]

RESEARCH NUMBER: USMA BS&L FALL 1999 Number 14
This research follows on RESEARCH NUMBER 5 conducted as a pilot study in Spring 1999.
RESEARCH TITLE: Evaluating Two Visual Interfaces to Mediated War Reserve Readiness Data
PRINCIPAL RESEARCHER / SUPERVISOR: MAJ Joseph Schafer / BS&L: Dr. Hah
PARTICIPANT'S TASK IN THE RESEARCH: Use two Java interfaces to answer representative tasks. Complete background and evaluative surveys.
INFO PERSON TO CANCEL: MAJ Schafer (914) 938-2407
joseph-schafer@usma.edu (dj4149) OFFICE: Thayer Hall room 113.
NOTE TO PARTICIPANTS: Please sign your name and e-mail address under the date and time you will participate in the research. Fill out your PL100 section number and instructor's name. Research participation has the priority of a scheduled class and lateness or absence will be treated accordingly. If you cannot attend, notify MAJ Schafer NLT 24hrs before.
MEET in Thayer Hall room 117 – (X-Lab) at the beginning of your scheduled hour.
Time Periods: Lesson 38-2 - Mon, 6 DEC 1999 K & L hours
Lesson 39-1 - Tues, 7 DEC 1999 A, B, E, F hours Lesson 39-2 - Wed, 8 DEC 1999 H, K, L hours

#	Date 1999	Time Lesson/Hour	Last Name, First Name	Section Number	Instructor	E-Mail #
1	6 Dec	1350-1445 38 / K				
2	6 Dec	1350-1445 38 / K				
3	6 Dec	1350-1445 38 / K				
4	6 Dec	1350-1445 38 / K				
5	6 Dec	1350-1445 38 / K				
6	6 Dec	1350-1445 38 / K				
7	6 Dec	1350-1445 38 / K				
8	6 Dec	1350-1445 38 / K				
9	6 Dec	1350-1445 38 / K				
10	6 Dec	1350-1445 38 / K				
11	6 Dec	1455-1550 38 / L				
12	6 Dec	1455-1550 38 / L				

MEET in TH117 – (X-Lab) at the beginning of your scheduled hour.
If you want to participate and none of these times work, email MAJ Schafer.

[2] Per Prof Hah, Fall 99, the experiment is very similar to the pilot studies, revalidation is not required.

G.2 Informed Consent Before Participation

Research Number: USMA BS&L Fall 1999 Number 14
(Office Use)

INFORMED CONSENT
(BEFORE PARTICIPATION)

I consent to participate in the research entitled:
Evaluating Two Visual Interfaces to Mediated War Reserve Readiness Data
conducted by the USMA Artificial Intelligence Office (Organization Name).

My task in the research is:
Log onto a web based interface evaluation tool. Use two Java interfaces to answer representative tasks. Complete background and evaluative surveys.

MAJ Schafer (Principal Researcher/Supervisor) or his/her representative explained the procedure and the expected duration of my participation. I am aware that although no physical or psychological harm is anticipated, I may withdraw from participating in this project at anytime, without penalty. I am also aware that I chose to participate in this research instead of taking a laboratory exercise. I was informed that after my participation, I will be briefed about the purpose of the research.

I acknowledge that my participation is free and voluntary. I understand the personal information I provide and the data collected will be used for research purposes only. They will be treated confidentially and will not be accessible to anyone outside the research team. A copy of this consent form will be supplied to me.

Date: _____

Printed Name _____ (Cadet)

Instructor's Name and Section Number _____ (Cadet)

Signed: _____ (Cadet)

Signed by: _____ (Experimenter/Data Collector)
Joseph Schafer, MAJ, USA

G.3 Informed Consent After Participation

Research Number: USMA BS&L Fall 1999 Number 14
(Office Use)

INFORMED CONSENT
(AFTER PARTICIPATION)

I have completed participation in the above research project. My participation lasted __one___ hour(s) and __zero___ minutes and I have been credited with ___one___ hour(s) of research time. The purpose of the research was

To compare and evaluate two different graphical user interfaces to agent-mediated army war reserve equipment readiness legacy data.

To assist planners and senior leaders in their ability to visualize the current readiness of AWR assets in the Afloat Set.

_____

_____ .

I was fully debriefed regarding the purpose of this project. I was also given the opportunity to obtain further information about the research. All my questions have been answered to my satisfaction.

Date: _____

Printed Name _____ (Cadet)

Signed: _____ (Cadet)

Signed by: _____ (Experimenter/Data Collector)
Joseph Schafer, MAJ, USA

G.4 Research Participation Worksheet

Research Number: USMA BS&L Fall 1999 Number 14

Participant's Name: _____ Date: _____

Participant's PL100 Instructor: _____Section/Hour: _____

Title of Research:

Evaluating Two Visual Interfaces to Mediated War Reserve Readiness Data

Name of Principal Researcher(s)/Supervisor(s): ___MAJ Joseph Schafer x2407_

At the conclusion of the participation in this study, the researcher(s) and the participant will, as a minimum, discuss the questions listed below. The participant will write the answers to the questions in the spaces provided. The participant will submit this worksheet to his/her PL100 instructor in order to receive credit for participation.

Questions:

1. What type of research method (experimental, non-experimental) was employed in this study?

Experimental

2. What kind of statistics (descriptive or inferential) and central tendency measures will be used?

Descriptive, T-Test,

3. What are the potential methodological problems in this research? (Hawthorne effect, demand characteristics, sensitivity of measures, self-fulfilling prophesy)

Turn In This Worksheet To Your PL100 Instructor As Soon As Possible.

APPENDIX H: EXAMPLE OF RAW DATA

This is one of 15 log files generated for subject L1 during the pilot study. The same format was used for the dissertation experiment.

The data has been modified from the original to fit this paper. Only changes in whitespace and font have been made.

```
Tue May 11 15:11:11 1999:
========
Client xlab01.eecs.usma.edu (socket 4)
connected...
LogComment. R:0.000102  U:0.000000  S:0.010000 :: CMT " rh1111 11-May-99 3:07:38 PM
VITAJIMI.class:start()"
LogComment. R:20.361246  U:0.000000  S:0.010000 :: CMT " x23671 11-May-99 3:07:59 PM Login
x23671login walley anthrax"
Login user.
LogComment. R:20.516834  U:0.000000  S:0.030000 :: CMT " x23671 11-May-99 3:07:59 PM User
'walley' authenticated."
Open mediator.
/home/fac/hermes/Data/Logistics/LIA.med
LogComment. R:20.794851  U:0.080000  S:0.040000 :: CMT " x23671 11-May-99 3:08:00 PM
Mediator file '/home/fac/hermes/Data/Logistics/LIA.med' open."
LogComment. R:70.397964  U:0.080000  S:0.040000 :: CMT " x23671 11-May-99 3:08:49 PM
Srvy$DEMG$DEM$01$b$$"
LogComment. R:78.855904  U:0.080000  S:0.040000 :: CMT " x23671 11-May-99 3:08:58 PM
Srvy$DEMG$DEM$02$c$$"
LogComment. R:85.457832  U:0.080000  S:0.040000 :: CMT " x23671 11-May-99 3:09:04 PM
Srvy$DEMG$DEM$03$b$$"
LogComment. R:94.616066  U:0.080000  S:0.050000 :: CMT " x23671 11-May-99 3:09:13 PM
Srvy$DEMG$DEM$04$b$$"
LogComment. R:107.708855  U:0.080000  S:0.050000 :: CMT " x23671 11-May-99 3:09:26 PM
Srvy$DEMG$DEM$05$b$$"
LogComment. R:112.743439  U:0.080000  S:0.060000 :: CMT " x23671 11-May-99 3:09:31 PM
Srvy$DEMG$DEM$06$b$$"
LogComment. R:118.423225  U:0.080000  S:0.060000 :: CMT " x23671 11-May-99 3:09:37 PM
Srvy$DEMG$DEM$07$b$$"
LogComment. R:130.637558  U:0.080000  S:0.060000 :: CMT " x23671 11-May-99 3:09:49 PM
Srvy$DEMG$DEM$08$b$$"
LogComment. R:136.850510  U:0.080000  S:0.060000 :: CMT " x23671 11-May-99 3:09:56 PM
Srvy$DEMG$DEM$09$b$$"
LogComment. R:139.384262  U:0.080000  S:0.070000 :: CMT " x23671 11-May-99 3:09:58 PM -----
VectCtrl.class:start() -- APS-3 *H: LOC Vector."
LogComment. R:139.806000  U:0.080000  S:0.070000 :: CMT " x23671 11-May-99 3:09:59 PM
VitaGUIstart "
LogComment. R:176.706207  U:0.080000  S:0.070000 :: CMT " x23671 11-May-99 3:10:35 PM
VectCtrl - Populate action"
LogComment. R:177.262131  U:0.080000  S:0.070000 :: CMT " x23671 11-May-99 3:10:36 PM
$$$$$$$$VPanel populate requested..."
LogComment. R:177.450867  U:0.080000  S:0.070000 :: CMT " x23671 11-May-99 3:10:36 PM
$$$$$$$$VPanel populate Data Found..."
LogComment. R:177.751312  U:0.080000  S:0.070000 :: CMT " x23671 11-May-99 3:10:36 PM
$$$$$$$$Vect_Ctrl-Query"
LogComment. R:178.318451  U:0.080000  S:0.070000 :: CMT " x23671 11-May-99 3:10:37 PM
```

APPENDIX I: SOURCE CODE

This appendix lists the perl script used to transform the raw log data from the format in Appendix H into the analysis format in Appendix J. This script illustrates how the raw data was analyzed for this research. It also describes the availability of the source code for the research systems developed for this work.

I.1 Data transform script

```
#!/usr/bin/perl #-w #- this prints warnings - but if a replace $16 does not exist, then get errors..
#finduid10.pl - this file starts with the -sort file and puts it out in a different format
# author:  joseph h Schafer, joseph-schafer@usma.edu, schaferj@seas.gwu.edu
# this script
#       gets the log files in a dir
#       finds the subject names in each log file
#       concatenates all the log files for each subject into the a file with the subject name
# finduid6.pl then creates a file with L.pfe - after fixlines
#  then fixcols creates -all.pfe  (finduid7 passes on only the lines of interest, not all
# then sortdata creates -sort.pfe which has all the data but is sorted
# then analdata only keeps records of interest and creates epoch seconds (finduid7 moved the
srvyoutfile to analdata now analfile is all task stuff)
# then fixtasksrvy puts out the final files
#  bext version will reduce the number of intermediate files and
#get us the info we need - total task time, user time (query proc time subtracted out)
# record if jimi or vita was first
# comment files, files for analysis, etc.
# need num correct, comments, totals

#$holdirs = $/;
undef $/; # not sure if needed, but undefines the input record separator - used by the filx lines subr
#$/ = $holdirs;

$Sect{DEMG} = 1;
$Sect{VITA} = 2;
$Sect{JIMI} = 3;
$Sect{COMP} = 4;
$Sect{JUIC} = 5;
$Sect{DEM} = 1;
$Sect{VIT} = 2;
$Sect{JIM} = 3;
$Sect{COM} = 4;
$Sect{JUI} = 5;


#task answers
%ans = qw(
        01 e
        02 a
        03 c
        04 b
        05 a
        06 d
        07 c
        08 c
        09 a
        10 b
        11 c
        12 b
);

#print "ans 01 is $ans{\"01\"} \n"; # this works

#confidence likert conversions
%conf = qw(
        a 1.0
        b .75
        c .50
        d .25
        e 0.0
```

```
);


#comparison survey likert conversions
%comp = qw(
        a 0.0
        b .25
        c .50
        d .75
        e 1.0
);

#print "begin finduid7.pl \n";

# finduid9.pl starts with the sorted file so these skipped
#getfilenames();
#getsubjects();
#writesubjectfile();
#fixlines();
#fixcols();
#sortdata();

$/="\n";  # if don't use this, then must use global, but LinNo not increment
getfilenamessort();
analdata();
fixtasksrvy();

#print "done";
#print " \n";

#
#-------subroutines
#

#========================
sub getfilenames {
        opendir(DOT,".") ||
                die "Cannot opendir . (serious dainbramage): $!";

        #get the filenames that end in .log
        @filenames = grep(/\.log$/, readdir(DOT));
        closedir(DOT);
}  #end getfilenames
#========================

#========================
sub getfilenamessort {

# main idea here is to use the current hash that we used before,
#but when we already have most of the data - don't really need to
#link up the filenames with the userids now.

        opendir(DOT,".") ||
                die "Cannot opendir . (serious dainbramage): $!";

        #get the filenames that end in .log
        @filenames = grep(/-sort\.pfe$/, readdir(DOT));
        #print @filenames;
        closedir(DOT);
        while ($filename = shift(@filenames)) {
               $_ = $filename;
                s/(.*)-sort\.pfe/$1/;
                $uid = $1;  #@uid works also

        #       print "subjects \n";
                #set up the subjects hash with filenames for each uid
                #$subjects {$uid} = $filename;  #this works, but want to add filename
                #Problem: For each key in a hash, only one scalar value is allowed, but you'd like to
use one key to store and retrieve multiple values.
                #That is, you'd like the value to be a list.
                #Solution: Use references to arrays as the hash values. Use push to append:

                push(@{ $subjects {$uid} }, $filename);
        }
}  #end getfilenamessort
#========================


#========================
sub getsubjects {
        #create a hash with a list of filenames
```

```
        while ($filename = shift(@filenames)) {
        open (FILE, $filename)||
                        die "can't open $filename: $!";;

                $_ = <FILE>;          # whole file now here
                s/.*Login (.*)login.*/$1/;
                $uid = $1;  #@uid works also

        #       print "subjects \n";
                #set up the subjects hash with filenames for each uid
                #$subjects {$uid} = $filename;  #this works, but want to add filename
                #Problem: For each key in a hash, only one scalar value is allowed, but you'd like to
use one key to store and retrieve multiple values.
                #That is, you'd like the value to be a list.
                #Solution: Use references to arrays as the hash values. Use push to append:

                push(@{ $subjects {$uid} }, $filename);
                #Then, dereference the value as an array reference when printing out the hash:
                #see printsubjects subroutine

        }
        close FILE;

} #end getsubjects
#========================


#========================
sub printsubjects {
        #Then, dereference the value as an array reference when printing out the hash:

        foreach $string (keys %subjects) {
                print "$string: @{$subjects{$string}}\n";
        }

        print "\n";
} #end printsubjects
#========================


#========================
sub writesubjectfile {

        foreach $string (keys %subjects) {
                $subjectfilename = $string . ".pfe";# dot is the concatenate operator
                open(SUBJECTFILE, ">>$subjectfilename")              or die "can't append to $string:
$!";
                binmode(SUBJECTFILE);

                foreach $logfilename (@{$subjects{$string}}){
                        #print "logfile: $logfilename";
                        open(LOGFILE, "< $logfilename")                                or die $!;
                        binmode(LOGFILE);
                        #write all the logfiles into subjectfile
                        print SUBJECTFILE <LOGFILE>;
                                #later may not write out intermediate files
                                #probably create another set of hashes with the data
                                #@ARRAY = <LOGFILE>;
                                # change ARRAY here
                        close (LOGFILE);
                        #print "\n";
                        #print "$string: @{$subjects{$string}}\n";
                }
        }

close(SUBJECTFILE)           or die "SUBJECTFILE didn't close: $!";

} #end writesubjectfile
#========================


#========================

# find duplicate words in paragraphs, possibly spanning line boundaries.
#   Use /x for space and comments, /i to match the both `is'
#   in "Is is this ok?", and use /g to find all dups.
# $/ = '';        # paragrep mode

# name          escape   hex      octal   decimal  control
```

```
# stty intr               \x03     \003     3        ^C
# stty eof                \x04     \004     4        ^D
# BEL            \a       \x07     \007     7        ^G
# BackSpace               \x08     \010     8        ^H
# Tab            \t       \x09     \011     9        ^I
# newline        \n       \x0a     \012     10       ^J
#
# FormFeed       \f       \x0c     \014     12       ^L
# return         \r       \x0d     \015     13       ^M
# space          " "      \x20     \040     32       " "
# windows if you do a \n (\x0a) then it comes out \r\n!
#
#lines of text normally end with
#        OS      End of Line
#        UNIX    \n
#        Mac     \r
#        DOS     \r\n

#the Win32 API) understands that / is a directory separator,
#so using / allows you to more easily port scripts between UNIX and Win32:
#open( MYFILE, '/temp/newfile.txt' );

#On win32 for text files, the \r\n characters are translated into \n when read
#from disk, and the ^Z character is read as an end-of-file marker.
#For binary files, no such translation is used.
# unfortunately in win converting 0a converts 0d0a!
##($infile, $outfile, $find, $replace, $options) = @ARGV; #take the files from the command line



#========================
sub fixdata {



} #end fixdata

#========================

#========================
sub fixlines {

        foreach $string (keys %subjects) {
                $subjectfilename = $string . ".pfe";
                open(SUBJECTFILE, "<$subjectfilename")              or die "can't append to $string:
$!";
                binmode(SUBJECTFILE);

                $outfilename = $string . "L.pfe"; # dot is the concatenate operator
                #print "$outfilename \n";

                open( OUTFILE, ">$outfilename" )|| die "Can't Open outfilename $outfilename";
                binmode( OUTFILE ); # crucial for binary files!

                #$_ = <SUBJECTFILE>;          # whole file now here
                # no difference between the $_ and the while
          while (<SUBJECTFILE>) { #try instead of $_ = <SUBJECTFILE>;
                s/\x0a/\x20/g;   #substitute space for \n globally on the current line
                s/\r/ /g; # substitute carrage return with space
                s/LogComment./\r\n./g; #was $buffer =~ s/LogComment./\r\nLogComment./g;
                s/(\w\w\w \w\w\w \d\d? \d\d?:\d\d:\d\d
\d\d\d\d:)/\r\n0.000000\t0.000000\t0.000000\t\t\t\t\t\t\t\t\t\t\t\t\|\t$1/g;
                s/(DBG:)/\r\n0.000000\t0.000000\t0.000000\t\t\t\t\t\t\t\t\t\t\t\|\t$1/g;
                s/\r\n\s*\r\n/\r\n/g;
                s/\.\s\sR:(\d+\.\d+)\s\sU:(\d+\.\d+)\s\sS:(\d+\.\d+) :: CMT /$1\t$2\t$3\t/g;
                # try to put the filename at the top
                #s/^\s*\n/\r\n0.000000\t0.000000\t0.000000\t\t\t\t\t\t\t\t\t\t\t\|\t$string\n/;
                print OUTFILE $_;
          } #end while
        } #end foreach
    close( SUBJECTFILE );
    close( OUTFILE );
}#end fixlines

#========================
sub fixcols {
        foreach $string (keys %subjects) {
               $subjectfilename = $string . "L.pfe";
                #print $subjectfilename;
                open(SUBJECTFILE, "<$subjectfilename")              or die "can't append to $string:
$!";
```

```
                binmode(SUBJECTFILE);
                $outfilename = $string . "-all.pfe"; # dot is the concatenate operator
                #print "$outfilename \n";
                open( OUTFILE, ">$outfilename" )|| die "Can't Open outfilename $outfilename";
                binmode( OUTFILE ); # crucial for binary files or files on multiple os with different
newline sequences!
                #$_ = <SUBJECTFILE>;          # whole file now here
                $dayhour = $string; # this is the subject name - used to be day hour and machine

                $/="\n";  # if don't use this, then must use global, but LinNo not increment
                        $LinNo =0;
                while (<SUBJECTFILE>) {

                        #email address get rid of space or $ before the subject name
                        #s/"( |\$)(\S+)/$2/; # yes that seems to work!  replace " and space or $ and
one or more non-white space with the non-white space
                        # final g not really neccessary now that doing line by line again with $/ set
to newline
                        s/"( |\$)(\S+)/$2/g; # yes that seems to work!  replace " and space or $ and
one or more non-white space with the non-white space
                        #date time
                #s/( |\$)(\d\d?-\w{3,10}-99) (\d\d?:\d\d:\d\d (A|P)M)( |\$)/\t$2\t$3\t\t\t\t\t\t\t'/;
                s/( |\$)(\d\d-\w\w\w-\d\d) (\d\d?:\d\d:\d\d (A|P)M)( |\$)/\t$2\t$3\t\t\t\t\t\t\t'/g;
                        #pad to move comments out - move task & srvy back below
                        #s/((A|P)M)\t\t\t\t\t\t\t'( |\$){1,2}(Srvy|Task)( |\$)([A-Z]{4})( |\$)([A-
Z]{3})( |\$)(\d\d)( |\$)(a|b|c|d|e)?( |\$)?(a|b|c|d|e)?(
|\$)?(.*)".*/$1\t$4\t$Sect{$6}\t$8\t$10\t$12\t$14\t$16/g;

                        #Problem with JIMI - file says TASK$JIMI$VIT instead of TASK$JIMI$JIM so pull
if from the JIMI
                        #s/((A|P)M)\t\t\t\t\t\t\t'( |\$){1,2}(Srvy|Task)( |\$)([A-Z]{4})( |\$)([A-
Z]{3})( |\$)(\d\d)( |\$)(a|b|c|d|e)?( |\$)?(a|b|c|d|e)?(
|\$)?(.*)".*/$1\t$4\t$Sect{$6}\t$8\t$10\t$12\t$14\t$16/g;
                        s/((A|P)M)\t\t\t\t\t\t\t'( |\$){1,2}(Srvy|Task)( |\$)([A-Z]{3})[AIGPC](
|\$)([A-Z]{3})( |\$)(\d\d)( |\$)(a|b|c|d|e)?( |\$)?(a|b|c|d|e)?(
|\$)?(.*)".*/$1\t$4\t$Sect{$6}\t$6\t$10\t$12\t$14\t$16/g;

                        s/\$*//g;  #kill any remaining $ delimiters
                        #substitute in above: ([A-Z]{4}) for (\w\w\w\w)
                        s/" */\t\|\t/g;  #replace the last " with a pipe
                        #number the lines
                        #s/(^|\n)/\n$LinNo\t$dayhour\t/g;  #add line number and uid to beginning -
since doing one big line look for \n
                        s/^/$LinNo\t$dayhour\t/;  #add line number and uid to beginning - since doing
one big line look for \n
                        $LinNo++;
                        #print "$LinNo \n";
                        #put in title line for fields if want to look at intermediate file in excel

        #s/^0\t$dayhour\t\s*\n/Lin\tDayHour\tRtime\tUtime\tStime\tEmail\tDate\tTime\tTest\tSect\tQcd\t
Qno\tAns\tConf\tComments\t\|\t\CodeCmt\n/;
        #
        s/^0\t$dayhour\t.*\n/Lin\tDayHour\tRtime\tUtime\tStime\tEmail\tDate\tTime\tTest\tSect\tQcd\tQn
o\tAns\tConf\tComments\t\|\t\CodeCmt\n/;
        #               s/\n/\t\|\n/;  # add a pipe after the last column not sure why
              if (/Task|Srvy|(Juice.(Jimi|Vita)GUIstart|Med_Ctrl-Execute_BTN_actionPerformed -
(|after)LaunchQueryFrame|Vect_Ctrl-(|after)Query)/) {
                        print OUTFILE $_ ;                #print the current line
              } #move this if here from anal data to reduce size of files
          } #end while
        } #end for each
        close( SUBJECTFILE );
        close( OUTFILE );
} #end fixcols



#========================
sub sortdata {
        foreach $string (keys %subjects) {
                $subjectfilename = $string . "-all.pfe";
                 #print $subjectfilename;
                 open(SUBJECTFILE, "<$subjectfilename")              or die "can't append to $string:
$!";
                 binmode(SUBJECTFILE);

                 $sortoutfilename = $string . "-sort.pfe"; # dot is the concatenate operator
                 open( SORTOUTFILE, ">$sortoutfilename" )|| die "Can't Open outfilename $outfilename";
                 binmode( SORTOUTFILE ); # crucial for binary files or files on multiple os with
different newline sequences!
```

```
                $dayhour = $string; # this is the subject name - used to be day hour and machine
                my @lines = ();

                while (<SUBJECTFILE>) {
                #move the analysis of the data to a new subr
                        /(^.*)(\d\d?):(\d\d):(\d\d )(.*)/;
                        #print "$2:$3:$4 ";

                        push @lines, {
                        BEGINNING        => $1,
                                HOUR             => $2,
                                MINUTE           => $3,
                                SECOND           => $4,
                                ENDING           => $5,
                        };
                } #end while

                        #print "sorted: \n";
                for my $line     (sort    {
                                $a->{HOUR}  <=> $b->{HOUR}
                                                ||
                                $a->{MINUTE}  <=> $b->{MINUTE}
                                                ||
                                $a->{SECOND}  <=> $b->{SECOND}
                                        } @lines
                                )
                        {
                                print SORTOUTFILE "$line->{BEGINNING}$line->{HOUR}:$line-
>{MINUTE}:$line->{SECOND}$line->{ENDING}\n";                #print the current line
                        } #end for my $line
        } #end for each
        close( SUBJECTFILE );
        close( SORTOUTFILE );
} #end sortdata


#=======================
sub analdata {
        foreach $string (keys %subjects) {
                $subjectfilename = $string . "-sort.pfe";
                #print $subjectfilename;
                open(SUBJECTFILE, "<$subjectfilename")              or die "can't append to $string:
$!";
                binmode(SUBJECTFILE);

                $sortoutfilename = $string . "-anal.pfe"; # dot is the concatenate operator
                open( ANALOUTFILE, ">$sortoutfilename" )|| die "Can't Open outfilename $outfilename";
                binmode( ANALOUTFILE ); # crucial for binary files or files on multiple os with
different newline sequences!

                $srvyoutfilename = $string . "-srvy.pfe"; # dot is the concatenate operator
                open( SRVYOUTFILE, ">$srvyoutfilename" )|| die "Can't Open outfilename $outfilename";
                binmode( SRVYOUTFILE ); # crucial for binary files or files on multiple os with
different newline sequences!

                $dayhour = $string; # this is the subject name - used to be day hour and machine
                my @lines = ();  #create local array
use Time::Local;
#$TIME = timelocal($sec, $min, $hours, $mday, $mon, $year);

            while (<SUBJECTFILE>) {

              #move this if which gets the lines of interest to the fixcols() above to reduce size of
intermediate files
              #if (/Task|Srvy|(Juice.(Jimi|Vita)GUIstart|Med_Ctrl-Execute_BTN_actionPerformed -
(|after)LaunchQueryFrame|Vect_Ctrl-(|after)Query)/) {
                # alernately (/Task/|/Srvy/|/(Juice.(Jimi|Vita)GUIstart|Med_Ctrl-
Execute_BTN_actionPerformed - (|after)LaunchQueryFrame|Vect_Ctrl-(|after)Query)/) {
                #trim off the front part of the data
                s/.*\t(\w+\t\d\d-\w\w\w-\d\d\t\d\d?:\d\d:\d\d (A|P)M.*)/$1/;
                s/(.*)(\t\|\W.*)/$1/;  # don't know what the prob was, but this fixed it(\t|\t\|\W.*)
                s/(.*(\d\d)-(\w\w\w)-(\d\d)\t(\d\d?):(\d\d):(\d\d))(.*)/$1$8/;
                $temptime=timelocal($7, $6, $5, $2, 11, $4); #timelocal($sec, $min, $hours, $mday,
$mon, $year)
                s/(.*)\W*/$1\t$temptime\n/;
                if (/Task/) {
                        #s/.*(\d\d?:\d\d:\d\d (A|P)M)\t(Task)\t(\d)\t([A-
Z]{3})\t(\d\d)\t(a|b|c|d|e)?\t?(a|b|c|d|e)?(.*)/$dayhour\t$1\t$3\t$4\t$5\t$6\t$7\t$8\t$9/;
                        print ANALOUTFILE $_ ;            #print the current line
                }
                elsif (/Srvy/) {
```

```
                        s/.*(Srvy)\t(\d)\t([A-
Z]{3})\t(\d\d)\t((a|b|c|d|e)|)\t*(.*)\W*\d\d\d\d\d\d\d\d\d\W*/$3\t$4\t$6$7\n/;
                        #\W*\d\d\d\d\d\d\d\d\d\W*
                        print SRVYOUTFILE $_ ;            #print the current line
                        print ANALOUTFILE $_ ;            #print the current line
                }
                elsif (/(Juice.(Jimi|Vita)GUIstart|Med_Ctrl-Execute_BTN_actionPerformed -
(|after)LaunchQueryFrame|Vect_Ctrl-(|after)Query)/) {
                        s/(.*)(Med_Ctrl-Execute_BTN_actionPerformed -
afterLaunchQueryFrame)/$1endquery/;
                        s/(.*)(Med_Ctrl-Execute_BTN_actionPerformed -
LaunchQueryFrame)/$1beginquery/;
                        s/(.*)Vect_Ctrl-afterQuery/$1endquery/;
                        s/(.*)Vect_Ctrl-Query/$1beginquery/;

                        print ANALOUTFILE $_ ;            #print the current line
                }

             # } #end if to just operate on the lines of interest -- moved to fixcols()
            } #end while
        } #end for each
        close( SUBJECTFILE );
        close( ANALOUTFILE );
        close( SRVYOUTFILE );
} #end analdata


#========================
sub fixtasksrvy {

                $allsvyfilename = "allsvyfile.pfe";
                open(ALLSVYFILE, ">$allsvyfilename")              or die "can't append to $string:
$!";
                binmode(ALLSVYFILE);
                #Demographic and coMpare
                print ALLSVYFILE
"Subj\td1\td2\tm1\tm2\tm3\tm4\tm5\tm6\tm7\tm8\tm9\tm10\tm11\tm12\tm13\tm14\tm15\tm16\tmv\tmj";

                $alltskfilename = "alltskfile.pfe";
                open(ALLTSKFILE, ">$alltskfilename")              or die "can't append to $string:
$!";
                binmode(ALLTSKFILE);
                print ALLTSKFILE
        "Subj\tOrd\tTreat\ta4\tc4\tu4\tn4\ta5\tc5\tu5\tn5\ta6\tc6\tu6\tn6\ta7\tc7\tu7\tn7\ta8\tc8\tu8\
tn8\ta9\tc9\tu9\tn9\ta10\tc10\tu10\tn10\ta11\tc11\tu11\tn11\ta12\tc12\tu12\tn12\taa\tac\tat\tau\tan\tc
c\tcu\tcn\ttreatuser\ta21\tc21\tu21\tn21\ta22\tc22\tu22\tn22\ta23\tc23\tu23\tn23\tordt\teas\thlp\tfst\
tinf\tgui\tlrn\tfun\tuse\tsrv\t";
                #\ta1\tc1\tu1\tn1\ta2\tc2\tu2\tn2\ta3\tc3\tu3\tn3  THE FIRST THREE ARE EXAMPLES
                open(COMMENTFILE, ">commentfile.pfe")              or die "can't append to $string:
$!";
                binmode(COMMENTFILE);


        foreach $string (keys %subjects) {
                $subjectfilename = $string . "-anal.pfe";
                print "$string ";
                open(SUBJECTFILE, "<$subjectfilename")              or die "can't append to $string:
$!";
                binmode(SUBJECTFILE);

                $taskoutfilename = $string . "-task.pfe"; # dot is the concatenate operator
                #print "$taskoutfilename \n";
                open( TASKOUTFILE, ">$taskoutfilename" )|| die "Can't Open outfilename $outfilename";
                binmode( TASKOUTFILE ); # crucial for binary files or files on multiple os with
different newline sequences!



                $dayhour = $string; # this is the subject name - used to be day hour and machine
                $firstgui = "";
                $stime =0; #start task
                $etime = 0; #end task
                $ttime = 0; #task total
                $qtime = 0; #query total
                $bqtime = 0; #beginquery
                $eqtime = 0; #endquery
                $utime = 0; #total user time for task with qtime subtacted
                $numq = 0; #number of queries
                $fixedj = 0;
                $fixedv = 0;
```

```
                print ALLSVYFILE "\n$string\t";
                $mv = 0.0;  #Total per subject of coMpare for Vitamin
                $mj = 0.0;  #total per subject of coMpare for Jimi
                @tmv ="";
                @tmj ="";

                #print ALLTSKFILE "\n$string\t";

        while (<SUBJECTFILE>) { #try instead of $_ = <SUBJECTFILE>;
                if (/(Vita|Jimi)GUIstart\W*(\d*)/) {
                        $stime = $2;
                        if ($firstgui eq "") {
                                #$firstgui = substr ($1, 0, 3);  #Needs to be 3 char & uppercase to
match
                                $firstgui = $1;  # do assignments instead of tests & changes for
speed
                                if ($firstgui eq "Jimi") {
                                        $firstgui = "JIM";
                                }
                                else {
                                        $firstgui = "VIT";
                                }
                        }
                        #print "test $firstgui $stime \n";
                }

                if (/beginquery\W*(\d*)/) {
                        $bqtime = $1;
                        $numq++;

                }
                if (/(.*)endquery\W*(\d*)/) {
                        $eqtime = $2;
                        if ($bqtime eq 0) {print "error $string $1\n";}
                        $qtime = $qtime + $eqtime - $bqtime;
                }

                if (/DEM|COM/) {
                        if (/\t(17|18|19)\t/) {
                                print COMMENTFILE "$string $_";
                        }
                        else {
                                s/(DEM)\t\d\d\t(\w)\W*/$2\t/;  #get rid of carriage return at end of
line to concatenate it all
                                s/(COM)\t(\d\d)\t(\w)\W*/$comp{$3}\t/;  #get rid of carriage return
at end of line to concatenate it all

                                $compsvy = $comp {$3};
                                if ($2 == 1)  { $mv += $compsvy; push (@tmv,"$compsvy\t"); }
                                if ($2 == 3)  { $mv += $compsvy; push (@tmv,"$compsvy\t"); }
                                if ($2 == 5)  { $mv += $compsvy; push (@tmv,"$compsvy\t"); }
                                if ($2 == 7)  { $mv += $compsvy; push (@tmv,"$compsvy\t"); }
                                if ($2 == 9)  { $mv += $compsvy; push (@tmv,"$compsvy\t"); }
                                if ($2 == 11) { $mv += $compsvy; push (@tmv,"$compsvy\t"); }
                                if ($2 == 13) { $mv += $compsvy; push (@tmv,"$compsvy\t"); }
                                if ($2 == 15) { $mv += $compsvy; push (@tmv,"$compsvy\t"); }
                                if ($2 == 2)  { $mj += $compsvy; push (@tmj,"$compsvy\t"); }
                                if ($2 == 4)  { $mj += $compsvy; push (@tmj,"$compsvy\t"); }
                                if ($2 == 6)  { $mj += $compsvy; push (@tmj,"$compsvy\t"); }
                                if ($2 == 8)  { $mj += $compsvy; push (@tmj,"$compsvy\t"); }
                                if ($2 == 10) { $mj += $compsvy; push (@tmj,"$compsvy\t"); }
                                if ($2 == 12) { $mj += $compsvy; push (@tmj,"$compsvy\t"); }
                                if ($2 == 14) { $mj += $compsvy; push (@tmj,"$compsvy\t"); }
                                if ($2 == 16) { $mj += $compsvy; push (@tmj,"$compsvy\t");

                                        #push(@{ $subjects {$uid} }, $filename);
                                        #push (@{$aft{$ttreat}},
"\n$string\t$order\t$ttreat\t$tans\t$tconf\t$utime\t$numq");

                                        #2800 rows...
                                        #foreach $gui (keys %aft) {
                                        #       print ALLTSKFILE "@{$aft{$gui}}";

                                        print ALLTSKFILE "@{$aft{\"VIT\"}} \t@tmv $mv";

                                        print ALLTSKFILE "@{$aft{\"JIM\"}} \t@tmj $mj";
                                        undef %aft;  #clear the hash

                                        #print "tmjv: @tmj @tmv";
                                        s/(.*)/$1$mv\t$mj/;
```

```
                                }
                                print ALLSVYFILE $_;
                        }
                }

                if (/Task.*(\d\d\d\d\d\d\d\d\d)/) {
                        $etime = $1;
                        #print $etime;
                        $ttime = $etime - $stime;
                        $utime = $ttime - $qtime;
                        #s/.*(\d\d?:\d\d:\d\d (A|P)M)\t(Task)\t(\d)\t([A-
Z]{3})\t(\d\d)\t(a|b|c|d|e)?\t?(a|b|c|d|e)?(.*)\W*\d\d\d\d\d\d\d\d\d\W*/$3\t$4\t$5\t$6\t$7\t$8\t$9\t$t
time\t$utime\t$numq\n/;
                        s/.*(\d\d?:\d\d:\d\d (A|P)M)\t(Task)\t(\d)\t([A-
Z]{3})\t(\d\d)\t(a|b|c|d|e)\t(a|b|c|d|e).*/$5\t$6\t$7\t$8\t$ttime\t$utime\t$numq/;

                        #Task variables
                        $ttreat=$5; #treatment VIT or JIM
                        $tques=$6;  #question number 01..12
                        $tempans=$7;     #Temporary answer a, b, c, d, e  -- 0 or 1 will be written
out based on $ans{}
                        $tconf=$conf{$8};  #confidence converted to 1..0 from abcde likert scale


                                if ($5 eq $firstgui) {
                                        $order = "1"; # this works also s/^/1/;
                                }
                                else {
                                        $order = "2"; # but not when expanded to ALLTSKFILE s/^/2/;
                                }

                        #the first task is number 02 (and so is the second) so fix it!
                        if (($fixedj == 0 )&& ($ttreat eq "JIM" )) {#fix first jimi task
                                s/JIM\t02/JIM\t01/;
                                $fixedj=1;
                                $tques="01";
                                #print "fixedj set $string\n";
                        }
                        if (($fixedv == 0) && ($ttreat eq "VIT")) {#fix first vita task
                                s/VIT\t02/VIT\t01/;
                                $fixedv=1;
                                $tques="01";
                                #print "fixedv set $string\n";
                        }
                        print TASKOUTFILE "$order$_";              #print the current line


                        #convert answer to 1 for correct, 0 for wrong with $ans hash
                        #had to wait until after the fixedv & fixedj routines!
                        if ($tempans eq $ans{$tques}) { $tans = 1;}
                        else {$tans = 0;}

                        #subject treatment totals
                        $stans += $tans;
                        $stconf += $tconf;
                        $stttime += $ttime;  #new in finduid10.pl
                        $stutime += $utime;
                        $stnumq += $numq;

                        #subject treatment correct only
                        $cstconf += $tconf*$tans;
                        $cstutime += $utime*$tans;
                        $cstnumq += $numq*$tans;


                        #actually the first three are just examples, so don't need to record for
analysis
                        #if ($tques == 1) {
                        #       print ALLTSKFILE
"\n$string\t$order\t$ttreat\t$tans\t$tconf\t$utime\t$numq";
                        #}
                        #if ($tques == 2) {
                        #       print ALLTSKFILE "\t$tans\t$tconf\t$utime\t$numq";
                        #}

                        #if ($tques == 3) {
                        #       print ALLTSKFILE "\t$tans\t$tconf\t$utime\t$numq";
                        #}


                        if ($tques == 4) {
```

```
                                #print ALLTSKFILE
"\n$string\t$order\t$ttreat\t$tans\t$tconf\t$utime\t$numq";
                                #Solution: Use references to arrays as the hash values. Use push to
append:
                                #push(@{ $subjects {$uid} }, $filename);
                                push (@{$aft{$ttreat}},
"\n$string\t$order\t$ttreat\t$tans\t$tconf\t$utime\t$numq");
                                #c21 are the correct numbers for the type 1  (21)  type 2 is 22, 3
is 23
                                $c21stans += $tans;
                                $c21stconf += $tconf*$tans;
                                $c21stutime += $utime*$tans;
                                $c21stnumq += $numq*$tans;

                        }

                        if ($tques == 5) {
                                push (@{$aft{$ttreat}},  "\t$tans\t$tconf\t$utime\t$numq");
                                $c21stans += $tans;
                                $c21stconf += $tconf*$tans;
                                $c21stutime += $utime*$tans;
                                $c21stnumq += $numq*$tans;
                        }
                        if ($tques == 6) {
                                push (@{$aft{$ttreat}},  "\t$tans\t$tconf\t$utime\t$numq");
                                $c21stans += $tans;
                                $c21stconf += $tconf*$tans;
                                $c21stutime += $utime*$tans;
                                $c21stnumq += $numq*$tans;
                        }

                        if ($tques == 7) {
                                push (@{$aft{$ttreat}},  "\t$tans\t$tconf\t$utime\t$numq");
                                $c22stans += $tans;
                                $c22stconf += $tconf*$tans;
                                $c22stutime += $utime*$tans;
                                $c22stnumq += $numq*$tans;
                        }
                        if ($tques == 8) {
                                push (@{$aft{$ttreat}},  "\t$tans\t$tconf\t$utime\t$numq");
                                $c22stans += $tans;
                                $c22stconf += $tconf*$tans;
                                $c22stutime += $utime*$tans;
                                $c22stnumq += $numq*$tans;
                        }

                        if ($tques == 9) {
                                push (@{$aft{$ttreat}},  "\t$tans\t$tconf\t$utime\t$numq");
                                $c22stans += $tans;
                                $c22stconf += $tconf*$tans;
                                $c22stutime += $utime*$tans;
                                $c22stnumq += $numq*$tans;
                        }
                        if ($tques == 10) {
                                push (@{$aft{$ttreat}},  "\t$tans\t$tconf\t$utime\t$numq");
                                $c23stans += $tans;
                                $c23stconf += $tconf*$tans;
                                $c23stutime += $utime*$tans;
                                $c23stnumq += $numq*$tans;
                        }

                        if ($tques == 11) {
                                push (@{$aft{$ttreat}},  "\t$tans\t$tconf\t$utime\t$numq");
                                $c23stans += $tans;
                                $c23stconf += $tconf*$tans;
                                $c23stutime += $utime*$tans;
                                $c23stnumq += $numq*$tans;
                        }
                        if ($tques == 12) {
                                push (@{$aft{$ttreat}},  "\t$tans\t$tconf\t$utime\t$numq");
                                $c23stans += $tans;
                                $c23stconf += $tconf*$tans;
                                $c23stutime += $utime*$tans;
                                $c23stnumq += $numq*$tans;
                                push (@{$aft{$ttreat}},
"\t$stans\t$stconf\t$stttime\t$stutime\t$stnumq");
                                push (@{$aft{$ttreat}},  "\t$cstconf\t$cstutime\t$cstnumq");
                                push (@{$aft{$ttreat}},  "\t$order$ttreat$string");
                                push (@{$aft{$ttreat}},
"\t$c21stans\t$c21stconf\t$c21stutime\t$c21stnumq");
```

```
                                    push (@{$aft{$ttreat}},
"\t$c22stans\t$c22stconf\t$c22stutime\t$c22stnumq");
                                    push (@{$aft{$ttreat}},
"\t$c23stans\t$c23stconf\t$c23stutime\t$c23stnumq");
                                    push (@{$aft{$ttreat}},  "\t$order$ttreat");
                                    #print "tans: $tans $tques $tempans $ans{$tques}. "

                                    #problem is that the @tmv & tmj are not prepared yet, so push
everything into an array and
                                    #print it from the srvy area!
                                    #if ($ttreat eq "VIT") {
                                    #        print ALLTSKFILE @tmv;
                                    #}
                                    #else {print ALLTSKFILE @tmj;}


                          #zero out subject treatment totals
                          $stans =0;
                          $stconf =0;
                          $stttime =0;
                          $stutime =0;
                          $stnumq =0;

                          #zero out subject treatment correct only
                          $cstconf =0;
                          $cstutime =0;
                          $cstnumq =0;

                          $c21stans=0;
                          $c21stconf =0;
                          $c21stutime =0;
                          $c21stnumq =0;
                          $c22stans =0;
                          $c22stconf =0;
                          $c22stutime =0;
                          $c22stnumq =0;
                          $c23stans =0;
                          $c23stconf =0;
                          $c23stutime =0;
                          $c23stnumq =0;

                          }
                          #zero out all the variables
                          $ttime=0; $stime=$etime; $etime=0; $qtime=0; $utime=0; $numq=0;
                          $bqtime = 0; $eqtime = 0;
                 }
            } #end while
          } #end for each
          close( SUBJECTFILE );
          close( TASKOUTFILE );
          close( ALLSVYFILE );
          close( ALLTSKFILE );
          close( COMMENTFILE);
} #end fixtasksrvy


sub bynumber{
#Problem: You want to sort a list of numbers, but Perl's sort (by default) sorts alphabetically in
ASCII order.
#Solution: Use Perl's sort function and the <=> numerical comparison operator:
#Example: @sorted = sort { $a <=> $b } @unsorted;

$a <=> $b;
} #end bynumber


sub by_number {
    if ($a < $b) {
        return -1;
    } elsif ($a == $b) {
        return 0;
    } elsif ($a > $b) {
        return 1;
    }
}
```

I.2 Developed Systems

Source code files for the systems developed and used for this research are much too large to include in this thesis. The source code for the research systems, VITAMIN, JIMI, and JUICE is available from http://www.seas.gwu.edu/~schaferj/. This site also links to information about the mediator-agent systems.

APPENDIX J: QUANTITATIVE DATA

This appendix includes tables for all of the quantitative data used for this research analysis.

The first table lists all of the variables with one variable in each row.

The next three tables list each experimental case, one per row with the variables across the top. The questions begin with number four because the first three questions for each treatment were examples that provided step-by-step solutions. These example questions provided an example of each of the three types of questions.

Name	Type	Label	Measure
subj	String	Subject	Nominal
ord	Numeric	Order	Ordinal
treat	String	Treatment	Nominal
a4	Numeric	AnswersQ4	Ordinal
c4	Numeric	ConfidenceQ4	Ordinal
u4	Numeric	UserTimeQ4	Scale
n4	Numeric	NumQueriesQ4	Ordinal
a5	Numeric	AnsweresQ5	Ordinal
c5	Numeric	ConfidenceQ5	Ordinal
u5	Numeric	UserTimeQ5	Scale
n5	Numeric	NumQueriesQ5	Ordinal
a6	Numeric	AnswersQ6	Ordinal
c6	Numeric	ConfidenceQ6	Ordinal
u6	Numeric	UserTimeQ6	Scale
n6	Numeric	NumQueriesQ6	Ordinal
a7	Numeric	AnswersQ7	Ordinal
c7	Numeric	ConfidenceQ7	Ordinal
u7	Numeric	UserTimeQ7	Scale
n7	Numeric	NumQueriesQ7	Ordinal
a8	Numeric	AnswersQ8	Ordinal
c8	Numeric	ConfidenceQ8	Ordinal
u8	Numeric	UserTimeQ8	Scale
n8	Numeric	NumQueriesQ8	Ordinal
a9	Numeric	AnswersQ9	Ordinal
c9	Numeric	ConfidenceQ9	Ordinal
u9	Numeric	UserTimeQ9	Scale
n9	Numeric	NumQueriesQ9	Ordinal
a10	Numeric	AnswersQ10	Ordinal
c10	Numeric	ConfidenceQ10	Ordinal
u10	Numeric	UserTimeQ10	Scale
n10	Numeric	NumQueriesQ10	Ordinal
a11	Numeric	AnswersQ11	Ordinal
c11	Numeric	ConfidenceQ11	Ordinal
u11	Numeric	UserTimeQ11	Scale
n11	Numeric	NumQueriesQ11	Ordinal
a12	Numeric	AnswersQ12	Ordinal
c12	Numeric	ConfidenceQ12	Ordinal
u12	Numeric	UserTimeQ12	Scale
n12	Numeric	NumQueriesQ12	Ordinal
aa	Numeric	AnswersAll	Ordinal
ac	Numeric	ConfideneAll	Scale
at	Numeric	TotalTimeAll	Scale
au	Numeric	UserTimeAll	Scale
an	Numeric	NumQueriesAll	Ordinal
cc	Numeric	ConfidenceCorrect	Scale
cu	Numeric	UserTimeCorrect	Scale
cn	Numeric	NumQueriesCorrect	Ordinal
mc	Numeric	ConfCorMean	Scale
mu	Numeric	UserTCorMean	Scale
mn	Numeric	NumQCorMean	Scale
ots	String	Order Treatment Subject aka treatuse	Nominal
a21	Numeric	AnswerT1	Ordinal
c21	Numeric	ConfidenceT1	Ordinal
u21	Numeric	UserTimeT1	Scale
n21	Numeric	NumQueriesT1	Ordinal
mc21	Numeric	ConfT1Mean	Scale
mu21	Numeric	UserT1Mean	Scale
mn21	Numeric	NumQT1Mean	Scale
a22	Numeric	AnswerT2	Ordinal
c22	Numeric	ConfidenceT2	Ordinal
u22	Numeric	UserTimeT2	Scale
n22	Numeric	NumQueriesT2	Ordinal
mc22	Numeric	ConfT2Mean	Scale
mu22	Numeric	UserT2Mean	Scale
mn22	Numeric	NumQT2Mean	Scale
a23	Numeric	AnswersT3	Ordinal
c23	Numeric	ConfidenceT3	Ordinal
u23	Numeric	UserTimeT3	Scale
n23	Numeric	NumQueriesT3	Ordinal
mc23	Numeric	ConfT3Mean	Scale
mu23	Numeric	UserT3Mean	Scale
mn23	Numeric	NumQT3Mean	Scale
ordt	String	Order and Treatment	Nominal
eas	Numeric	Ease of Use	Ordinal
hlp	Numeric	Helpful	Ordinal
fst	Numeric	Quickly	Ordinal
inf	Numeric	Information Quality	Ordinal
gui	Numeric	Interface Quality	Ordinal
lrn	Numeric	Learn About Data	Ordinal
fun	Numeric	Enjoyable	Ordinal
use	Numeric	Usable	Ordinal
srv	Numeric	Usability Survey Sum	Scale

Table 36 Variables

SUBJ	ORD	TREAT	A4	C4	U4	N4	A5	C5	U5	N5	A6	C6	U6	N6	A7	C7	U7	N7	A8	C8	U8	N8	A9	C9	U9	N9	A10	C10	U10	N10	A11	C11	U11	N11	A12	C12	U12	N12
2201	1	JIM	0	.00	92	2	1	.50	33	1	0	.00	34	0	1	.50	63	1	1	.50	50	1	1	.50	102	2	1	.50	46	1	0	.00	93	1	1	.00	14	0
2201	2	VIT	0	.50	40	0	1	.50	29	1	0	.50	15	0	1	.75	28	1	1	.75	20	1	1	.75	37	1	1	.75	72	4	1	1.00	41	2	1	1.00	45	2
2598	2	JIM	1	1.00	49	2	1	1.00	66	2	1	1.00	19	0	1	1.00	231	3	1	1.00	154	3	1	1.00	75	1	1	1.00	47	1	1	1.00	78	1	1	1.00	61	1
2598	1	VIT	1	1.00	18	0	1	1.00	14	0	1	1.00	11	0	1	1.00	50	2	1	1.00	17	1	1	1.00	23	1	1	1.00	92	6	1	1.00	41	3	1	1.00	57	2
2961	1	JIM	1	1.00	270	2	1	1.00	30	0	1	1.00	43	0	1	1.00	173	1	1	1.00	60	1	1	1.00	52	1	1	1.00	66	1	1	1.00	101	1	1	1.00	51	1
2961	2	VIT	1	1.00	30	0	1	1.00	15	1	1	1.00	16	0	1	1.00	29	1	1	1.00	18	1	1	1.00	16	1	1	1.00	61	5	1	1.00	28	1	1	1.00	27	1
3749	2	JIM	1	1.00	26	1	1	1.00	15	0	1	1.00	15	0	1	1.00	92	1	1	1.00	46	1	1	1.00	40	1	1	1.00	38	1	1	1.00	73	1	1	1.00	45	1
3749	1	VIT	1	1.00	58	0	1	.75	45	1	1	1.00	12	0	1	1.00	18	1	1	1.00	16	1	1	1.00	14	1	1	1.00	60	5	1	1.00	33	1	1	1.00	30	2
3779	2	JIM	1	1.00	40	1	1	1.00	68	1	0	.00	192	3	1	1.00	74	1	1	1.00	51	1	1	1.00	62	1	1	1.00	42	1	1	1.00	122	2	1	1.00	55	1
3779	1	VIT	1	1.00	52	0	1	1.00	22	1	0	1.00	22	0	1	1.00	41	2	1	1.00	33	2	1	1.00	26	2	1	1.00	82	6	1	1.00	18	1	1	1.00	43	1
3793	1	JIM	1	1.00	227	6	1	.75	48	1	1	1.00	31	0	1	1.00	127	2	1	1.00	318	3	1	1.00	95	1	1	1.00	38	1	1	.75	79	1	1	.75	55	1
3793	2	VIT	1	1.00	44	1	1	1.00	13	0	1	1.00	19	0	1	1.00	21	1	1	1.00	25	1	1	1.00	18	1	1	.75	97	7	1	1.00	25	1	1	1.00	39	1
3795	1	JIM	1	1.00	116	4	1	1.00	81	2	1	.75	98	1	1	1.00	192	3	1	1.00	110	3	1	.75	69	1	1	1.00	45	2	1	.75	56	1	1	1.00	49	1
3795	2	VIT	1	1.00	13	0	1	1.00	10	0	1	1.00	18	0	1	1.00	20	1	1	1.00	25	2	1	1.00	21	2	1	1.00	32	5	1	1.00	21	1	1	1.00	18	1
5161	2	JIM	1	1.00	208	2	1	1.00	38	1	1	1.00	20	0	1	1.00	105	2	1	1.00	50	1	1	1.00	109	2	1	.25	131	3	1	.75	142	2	1	.75	61	1
5161	1	VIT	1	1.00	24	0	1	1.00	14	1	1	1.00	17	0	1	1.00	43	1	1	1.00	34	1	1	1.00	50	1	1	1.00	66	5	1	1.00	24	1	1	1.00	36	1
5407	2	JIM	0	.00	113	2	1	.25	192	5	1	.50	63	1	0	.00	147	6	0	.00	117	4	1	.00	69	3	1	1.00	78	1	0	.00	114	2	1	.00	211	3
5407	1	VIT	1	.75	24	0	1	.50	26	1	1	.75	41	1	1	.50	53	2	1	.25	39	2	1	.50	44	1	1	.75	60	7	1	.75	42	3	1	.75	21	2
5424	1	JIM	1	1.00	107	2	1	1.00	15	0	1	1.00	86	0	1	1.00	143	1	1	1.00	63	1	0	1.00	112	1	1	1.00	98	1	1	1.00	227	3	1	1.00	85	1
5424	2	VIT	1	1.00	35	3	1	1.00	32	1	1	1.00	13	0	1	1.00	57	2	1	1.00	25	1	1	1.00	45	1	1	1.00	103	5	1	1.00	52	2	1	1.00	43	2
5516	1	JIM	1	.75	85	1	1	.75	27	0	1	.75	42	0	1	.75	204	3	1	.75	58	1	1	.75	160	1	1	.75	78	1	0	.25	415	4	1	.50	127	1
5516	2	VIT	1	1.00	23	0	1	1.00	16	0	1	1.00	13	0	1	1.00	33	2	1	1.00	31	3	0	1.00	26	1	1	1.00	59	4	1	.75	31	1	1	1.00	40	1
5546	2	JIM	1	1.00	55	1	1	.75	16	0	1	1.00	27	0	1	1.00	156	3	1	1.00	33	1	1	1.00	59	1	0	1.00	62	1	1	1.00	144	2	1	1.00	50	1
5546	1	VIT	1	.75	57	0	1	1.00	60	0	1	1.00	24	0	1	1.00	32	1	1	1.00	14	1	1	1.00	22	1	0	1.00	45	6	1	1.00	41	2	1	1.00	28	2
5582	2	JIM	1	1.00	54	1	1	1.00	21	0	1	1.00	23	0	1	1.00	126	2	1	1.00	59	1	1	1.00	54	1	1	1.00	155	3	1	1.00	100	2	1	1.00	55	1
5582	1	VIT	1	1.00	23	0	1	1.00	19	0	1	1.00	31	0	1	1.00	34	1	1	1.00	28	1	1	1.00	32	1	1	1.00	88	4	1	.75	84	1	1	1.00	51	1
5587	2	JIM	1	1.00	43	3	1	1.00	14	0	1	1.00	22	0	1	1.00	151	1	1	1.00	84	1	1	.75	89	1	1	1.00	120	6	0	.25	78	1	0	.00	26	0
5587	1	VIT	1	1.00	32	0	1	1.00	20	0	1	1.00	14	0	1	1.00	55	2	1	1.00	17	1	1	1.00	54	1	1	1.00	83	9	1	1.00	27	1	1	1.00	38	2
5602	2	JIM	1	1.00	44	1	1	1.00	30	0	1	1.00	37	1	1	1.00	159	2	1	1.00	50	1	1	1.00	74	1	1	1.00	47	1	1	1.00	70	1	1	1.00	57	1
5602	1	VIT	1	1.00	26	0	1	1.00	15	0	1	1.00	10	0	1	1.00	41	2	1	1.00	18	1	1	1.00	21	1	1	1.00	80	6	1	1.00	30	1	1	1.00	23	1
5607	2	JIM	1	1.00	49	1	1	1.00	34	1	1	1.00	27	0	1	1.00	96	1	1	1.00	53	1	1	1.00	65	1	1	1.00	54	1	1	1.00	128	1	0	.75	5	0
5607	1	VIT	0	1.00	23	0	1	1.00	19	1	1	1.00	23	0	1	1.00	31	1	1	1.00	62	1	1	1.00	18	1	1	1.00	61	3	1	1.00	34	1	1	1.00	20	1
5621	1	JIM	1	.75	53	1	1	.75	17	0	1	1.00	25	0	1	.75	126	3	1	.75	64	1	1	1.00	44	1	1	1.00	37	1	1	.50	152	1	0	.75	79	1
5621	2	VIT	1	1.00	21	0	1	.75	15	0	1	1.00	10	0	1	1.00	25	1	1	1.00	18	1	1	1.00	18	1	1	.75	52	5	1	.75	31	2	1	.75	34	2
5632	1	JIM	1	1.00	83	4	1	1.00	127	3	0	1.00	86	3	1	.75	208	3	1	1.00	98	2	1	1.00	50	1	1	1.00	54	1	1	1.00	102	2	1	1.00	64	1
5632	2	VIT	1	1.00	14	0	1	1.00	14	0	1	1.00	27	0	1	1.00	34	2	1	1.00	23	2	1	1.00	17	1	1	1.00	68	5	1	1.00	29	2	1	1.00	27	1
5642	1	JIM	1	1.00	43	3	1	1.00	25	0	1	.75	51	0	1	1.00	91	1	1	1.00	63	1	1	1.00	67	1	1	1.00	42	1	1	1.00	256	4	1	1.00	67	1
5642	2	VIT	1	1.00	19	0	1	1.00	27	1	1	1.00	18	0	1	1.00	39	2	1	1.00	21	1	1	1.00	30	1	1	1.00	83	5	1	1.00	24	2	1	1.00	25	2
5649	2	JIM	1	1.00	92	1	1	1.00	23	0	1	1.00	31	0	1	1.00	106	1	1	1.00	68	1	1	1.00	54	1	1	1.00	53	1	1	1.00	68	1	1	1.00	102	1
5649	1	VIT	1	1.00	25	0	1	1.00	13	0	1	1.00	14	0	1	1.00	54	2	1	1.00	30	1	1	.75	75	1	1	1.00	72	5	1	1.00	51	1	1	1.00	34	1
5679	1	JIM	1	.75	82	1	1	.75	22	0	1	.75	25	0	1	1.00	167	5	1	1.00	55	1	1	1.00	68	1	1	1.00	40	1	1	.75	132	2	1	.75	52	1
5679	2	VIT	1	1.00	25	0	1	.75	36	1	1	1.00	16	0	1	1.00	72	2	1	1.00	34	2	1	1.00	35	2	0	1.00	81	9	1	1.00	24	2	1	1.00	29	2
5687	1	JIM	1	.50	85	2	1	.75	51	1	1	.75	74	3	0	.50	77	1	1	.75	46	1	1	1.00	45	1	1	1.00	60	2	1	1.00	128	2	1	1.00	54	1
5687	2	VIT	1	1.00	30	0	1	1.00	13	0	1	1.00	10	0	1	1.00	19	1	1	1.00	15	1	0	1.00	12	1	1	1.00	36	5	1	1.00	19	1	1	1.00	44	3
5698	1	JIM	1	1.00	57	1	1	1.00	21	0	1	1.00	48	0	1	1.00	218	2	0	.25	98	1	1	1.00	82	2	1	1.00	39	1	1	1.00	77	1	1	1.00	73	1
5698	2	VIT	1	.75	43	1	1	1.00	34	1	1	1.00	12	0	1	1.00	35	1	1	1.00	20	1	1	1.00	19	1	1	1.00	54	6	1	1.00	20	2	1	1.00	43	3
5699	1	JIM	1	1.00	94	1	1	1.00	229	4	1	1.00	66	1	1	1.00	112	2	1	1.00	65	1	1	.75	62	1	1	1.00	142	3	0	.00	165	4	0	.50	158	2
5699	2	VIT	1	1.00	25	0	1	.75	42	2	1	1.00	19	0	1	1.00	21	1	1	1.00	20	1	1	1.00	14	1	1	.75	42	6	1	1.00	34	1	1	1.00	37	1
5715	2	JIM	1	1.00	42	1	1	.50	187	1	1	1.00	99	1	1	1.00	204	4	1	1.00	56	1	1	.75	106	1	1	1.00	110	3	1	1.00	98	1	1	1.00	49	1
5715	1	VIT	1	1.00	72	3	1	.75	41	1	1	1.00	36	0	1	1.00	34	0	1	1.00	27	1	0	.50	61	1	1	1.00	86	5	1	1.00	35	2	1	1.00	35	2
5736	1	JIM	1	.75	68	2	1	.75	16	0	1	.50	24	0	1	.75	124	1	1	.75	109	1	1	.75	126	2	1	.75	47	1	1	.75	70	1	1	.75	68	1
5736	2	VIT	0	1.00	34	1	1	1.00	9	0	1	.75	9	0	1	1.00	16	1	1	1.00	27	1	1	1.00	18	1	1	.75	71	5	1	1.00	41	1	1	1.00	30	1
5737	1	JIM	1	1.00	171	1	1	1.00	14	0	1	1.00	35	0	1	1.00	222	3	1	1.00	67	1	1	1.00	193	1	1	1.00	232	4	0	1.00	169	1	0	1.00	76	1
5737	2	VIT	1	1.00	33	0	1	1.00	16	0	1	1.00	20	0	1	1.00	32	1	1	1.00	21	1	1	1.00	26	1	1	1.00	94	6	1	1.00	38	2	1	1.00	44	2
5754	2	JIM	1	1.00	130	2	1	.75	24	0	1	1.00	22	0	1	.75	153	2	1	1.00	56	1	0	1.00	48	1	1	1.00	76	1	0	.25	190	5	0	.25	66	1
5754	1	VIT	1	1.00	29	0	1	1.00	36	1	1	1.00	23	0	1	1.00	25	1	1	1.00	24	1	1	1.00	21	1	1	1.00	98	5	1	1.00	30	1	1	1.00	30	1
5778	2	JIM	1	1.00	116	1	1	1.00	17	0	1	1.00	24	0	1	.75	125	2	1	.75	58	1	1	.75	60	1	1	1.00	54	1	0	.75	65	1	0	.75	97	2
5778	1	VIT	1	1.00	83	0	1	1.00	44	1	1	1.00	33	0	1	1.00	25	1	1	.75	16	1	1	1.00	24	1	1	1.00	86	6	1	1.00	44	2	1	1.00	50	2
5801	1	JIM	1	1.00	129	2	1	1.00	66	1	1	1.00	47	0	1	1.00	107	1	1	1.00	82	2	1	1.00	99	1	1	1.00	58	1	1	1.00	369	6	1	1.00	63	1
5801	2	VIT	1	1.00	26	0	1	1.00	48	1	1	1.00	23	0	1	1.00	66	2	1	1.00	24	1	1	1.00	34	1	1	1.00	55	4	1	1.00	30	1	1	1.00	23	1
5810	1	JIM	0	.75	68	1	1	1.00	107	2	0	.25	81	0	1	1.00	132	1	0	.00	242	5	1	1.00	74	1	1	.50	61	1	1	.00	193	3	0	1.00	128	2
5810	2	VIT	1	1.00	25	0	1	1.00	10	0	1	1.00	12	0	1	1.00	41	3	1	1.00	27	1	1	1.00	29	1	1	1.00	100	6	1	1.00	32	1	1	1.00	31	1
5812	2	JIM	1	1.00	91	1	1	1.00	19	0	1	1.00	26	0	1	1.00	208	5	1	1.00	54	1	1	1.00	72	1	1	1.00	65	1	1	1.00	203	3	1	1.00	48	1
5812	1	VIT	1	1.00	132	1	1	1.00	32	0	1	1.00	42	0	1	1.00	31	1	1	1.00	28	1	1	1.00	27	0	1	1.00	58	2	1	1.00	37	1	1	1.00	30	1
5816	2	JIM	1	1.00	49	1	1	.75	18	0	1	.75	24	0	1	.25	243	2	1	.25	78	1	1	.25	64	1	1	.25	56	1	1	.50	145	3	1	.50	159	3
5816	1	VIT	1	.50	90	0	1	.50	21	1	1	.50	27	0	1	1.00	24	1	1	1.00	26	1	1	1.00	24	1	1	1.00	64	6	1	.75	31	2	1	.75	36	2
5836	1	JIM	0	.00	134	3	1	1.00	78	1	0	.00	77	1	0	.00	67	1	1	1.00	62	1	1	.75	63	1	1	1.00	54	1	1	.75	43	1	0	.25	89	2
5836	2	VIT	0	.00	16	1	1	.50	39	1	0	.00	12	0	1	.00	28	0	0	.25	71	2	0	1.00	45	4	1	.75	20	1	1	1.00	42	3	1	1.00	37	2
5856	1	JIM	1	.75	130	4	1	1.00	43	1	1	1.00	82	2	1	.75	82	1	1	.75	60	1	1	.75	60	2	1	.75	57	1	1	.75	234	3	1	.50	57	1
5856	2	VIT	1	1.00	19	1	1	1.00	13	0	1	1.00	49	1	1	1.00	20	1	1	1.00	17	1	1	1.00	14	1	1	1.00	61	5	1	1.00	20	2	1	.75	46	3
5873	2	JIM	1	1.00	33	1	1	1.00	141	3	1	1.00	25	0	1	1.00	98	2	1	1.00	61	1	0	1.00	85	2	1	1.00	88	2	1	1.00	87	1	1	1.00	62	1
5873	1	VIT	1	1.00	27	0	1	1.00	55	2	1	1.00	19	0	1	1.00	36	1	1	1.00	26	1	1	1.00	22	1	0	1.00	91	6	1	1.00	33	1	1	1.00	31	1
5875	2	JIM	1	.75	67	1	1	.75	31	1	1	.75	29	0	1	.75	180	4	1	.75	56	1	1	.75	95	1	1	.75	57	1	1	.25	155	3	1	.50	72	1
5875	1	VIT	1	.50	142	1	1	.75	22	0	1	1.00	43	1	1	.75	40	1	1	.75	27	1	1	.75	90	5	1	.75	100	5	1	.50	39	1	1	.75	76	1
5886	1	JIM	1	.25	23	0	1	1.00	36	1	1	.75	78	1	0	.25	125	1	1	1.00	71	2	0	.75	56	1	1	1.00	53	1	0	.50	89	1	0	.75	104	2
5886	2	VIT	1	1.00	17	0	1	1.00	10	0	1	1.00	14	0	1	1.00	17	1	1	1.00	16	1	1	.75	30	1	1	1.00	46	4	1	1.00	21	1	1	1.00	27	1
5892	1	JIM	0	1.00	104	1	1	1.00	21	0	1	1.00	33	0	1	.75	132	1	1	1.00	55	1	1	1.00	83	1	1	1.00	71	1	1	.50	160	1	1	.50	179	1
5892	2	VIT	1	1.00	56	2	1	.75	14	0	1	1.00	15	0	1	1.00	26	1	1	1.00	32	2	1	1.00	16	1	1	1.00	57	5	1	1.00	42	1	1	1.00	37	2
5893	1	JIM	0	.25	111	1	1	.75	64	1	1	.50	24	0	1	.00	129	4	1	.25	142	2	1	.50	80	1	0	.00	42	1	1	.00	83	2	0	.25	64	0
5893	2	VIT	1	.50	21	0	1	.50	10	0	1	.25	21	0	1	.50	44	1	1	.50	40	1	1	.75	39	1	1	.75	123	4	1	.25	47	1	0	.25	40	0
5898	1	JIM	1	1.00	53	2	1	1.00	41	1	1	1.00	40	1	1	1.00	53	1	1	1.00	35	1	1	1.00	36	1	1	1.00	37	1	1	1.00	139	2	1	1.00	31	1
5898	2	VIT	1	1.00	19	0	1	1.00	9	0	1	1.00	10	0	1	1.00	17	1	1	1.00	13	1	1	1.00	17	1	1	.75	49	6	1	.75	23	1	1	1.00	29	1
5902	2	JIM	1	.75	151	7	1	1.00	20	1	0	.75	242	8	1	1.00	54	3	1	1.00	25	1	0	1.00	43	1	1	1.00	102	4	0	.50	74	0	0	.00	2	0
5902	1	VIT	0	1.00	73	0	1	1.00	30	1	1	1.00	13	0	1	1.00	31	1	1	1.00	22	1	1	1.00	25	3	1	1.00	61	5	1	1.00	24	1	1	1.00	37	1

Table 37 Subject Task Responses

SUBJ	AA	AC	AT	AU	AN	CC	CU	CN	MC	MU	MN	OTS	A21	C21	U21	N21	MC21	MU21	MN21	A22	C22	U22	N22	MC22	MU22	MN22	A23	C23	U23	N23	MC23	MU23	MN23
2201	6.0	2.50	1011	527	9	2.50	308	6	.42	51.33	1.00	1j2201	1	.50	33	1	.50	33.00	1.00	3	1.50	215	4	.50	71.67	1.33	2	.50	60	1	.25	30.00	.50
2201	7.0	6.50	668	327	12	5.50	272	12	.79	38.86	1.71	2v2201	1	.50	29	1	.50	29.00	1.00	3	2.25	85	3	.75	28.33	1.00	3	2.75	158	8	.92	52.67	2.67
2598	9.0	9.00	1171	780	14	9.00	780	14	1.00	86.67	1.56	2j2598	3	3.00	134	4	1.00	44.67	1.33	3	3.00	460	7	1.00	153.33	2.33	3	3.00	186	3	1.00	62.00	1.00
2598	9.0	9.00	657	323	15	9.00	323	15	1.00	35.89	1.67	1v2598	3	3.00	43	0	1.00	14.33	.00	3	3.00	90	4	1.00	30.00	1.33	3	3.00	190	11	1.00	63.33	3.67
2961	9.0	9.00	1419	846	8	9.00	846	8	1.00	94.00	.89	1j2961	3	3.00	343	2	1.00	114.33	.67	3	3.00	285	3	1.00	95.00	1.00	3	3.00	218	3	1.00	72.67	1.00
2961	9.0	9.00	450	240	11	9.00	240	11	1.00	26.67	1.22	2v2961	3	3.00	61	1	1.00	20.33	.33	3	3.00	63	3	1.00	21.00	1.00	3	3.00	116	7	1.00	38.67	2.33
3749	9.0	9.00	736	390	7	9.00	390	7	1.00	43.33	.78	2j3749	3	3.00	56	1	1.00	18.67	.33	3	3.00	178	3	1.00	59.33	1.00	3	3.00	156	3	1.00	52.00	1.00
3749	9.0	8.75	560	286	12	8.75	286	12	.97	31.78	1.33	1v3749	3	2.75	115	1	.92	38.33	.33	3	3.00	48	3	1.00	16.00	1.00	3	3.00	123	8	1.00	41.00	2.67
3779	8.0	8.00	1201	706	12	8.00	514	9	1.00	64.25	1.13	2j3779	2	2.00	108	2	1.00	54.00	1.00	3	3.00	187	3	1.00	62.33	1.00	3	3.00	219	4	1.00	73.00	1.33
3779	8.0	9.00	637	339	15	8.00	317	15	1.00	39.63	1.88	1v3779	2	2.00	74	1	1.00	37.00	.50	3	3.00	100	6	1.00	33.33	2.00	3	3.00	143	8	1.00	47.67	2.67
3793	9.0	8.25	1604	1018	16	8.25	1018	16	.92	113.11	1.78	1j3793	3	2.75	306	7	.92	102.00	2.33	3	3.00	540	6	1.00	180.00	2.00	3	2.50	172	3	.83	57.33	1.00
3793	9.0	8.75	613	301	13	8.75	301	13	.97	33.44	1.44	2v3793	3	3.00	76	1	1.00	25.33	.33	3	3.00	64	3	1.00	21.33	1.00	3	2.75	161	9	.92	53.67	3.00
3795	9.0	8.25	1538	816	18	8.25	816	18	.92	90.67	2.00	1j3795	3	2.75	295	7	.92	98.33	2.33	3	2.75	371	7	.92	123.67	2.33	3	2.75	150	4	.92	50.00	1.33
3795	9.0	9.00	369	178	12	9.00	178	12	1.00	19.78	1.33	2v3795	3	3.00	41	0	1.00	13.67	.00	3	3.00	66	5	1.00	22.00	1.67	3	3.00	71	7	1.00	23.67	2.33
5161	9.0	7.75	1309	864	14	7.75	864	14	.86	96.00	1.56	2j5161	3	3.00	266	3	1.00	88.67	1.00	3	3.00	264	5	1.00	88.00	1.67	3	1.75	334	6	.58	111.33	2.00
5161	9.0	9.00	715	308	11	9.00	308	11	1.00	34.22	1.22	1v5161	3	3.00	55	1	1.00	18.33	.33	3	3.00	127	3	1.00	42.33	1.00	3	3.00	126	7	1.00	42.00	2.33
5407	5.0	1.75	1608	1104	27	1.75	613	13	.35	122.60	2.60	2j5407	2	.75	255	6	.38	127.50	3.00	1	.00	69	3	.00	69.00	3.00	2	1.00	289	4	.50	144.50	2.00
5407	9.0	5.50	700	350	19	5.50	350	19	.61	38.89	2.11	1v5407	3	2.00	91	2	.67	30.33	.67	3	1.25	136	5	.42	45.33	1.67	3	2.25	123	12	.75	41.00	4.00
5424	8.0	9.00	1574	936	10	8.00	824	9	1.00	103.00	1.13	1j5424	3	3.00	208	2	1.00	69.33	.67	2	2.00	206	2	1.00	103.00	1.00	3	3.00	410	5	1.00	136.67	1.67
5424	9.0	9.00	707	405	17	9.00	405	17	1.00	45.00	1.89	2v5424	3	3.00	80	4	1.00	26.67	1.33	3	3.00	127	4	1.00	42.33	1.33	3	3.00	198	9	1.00	66.00	3.00
5516	8.0	6.00	1892	1196	12	5.75	781	8	.72	97.63	1.00	1j5516	3	2.25	154	1	.75	51.33	.33	3	2.25	422	5	.75	140.67	1.67	2	1.25	205	2	.63	102.50	1.00
5516	8.0	8.75	520	272	12	7.75	246	11	.97	30.75	1.38	2v5516	3	3.00	52	0	1.00	17.33	.00	2	2.00	64	5	1.00	32.00	2.50	3	2.75	130	6	.92	43.33	2.00
5546	8.0	8.75	1037	602	10	7.75	540	9	.97	67.50	1.13	2j5546	3	2.75	98	1	.92	32.67	.33	3	3.00	248	5	1.00	82.67	1.67	2	2.00	194	3	1.00	97.00	1.50
5546	8.0	8.75	710	323	13	7.75	278	7	.97	34.75	.88	1v5546	3	2.75	141	0	.92	47.00	.00	3	3.00	68	3	1.00	22.67	1.00	2	2.00	69	4	1.00	34.50	2.00
5582	9.0	9.00	1201	647	11	9.00	647	11	1.00	71.89	1.22	2j5582	3	3.00	98	1	1.00	32.67	.33	3	3.00	239	4	1.00	79.67	1.33	3	3.00	310	6	1.00	103.33	2.00
5582	9.0	8.75	706	390	9	8.75	390	9	.97	43.33	1.00	1v5582	3	3.00	73	0	1.00	24.33	.00	3	3.00	94	3	1.00	31.33	1.00	3	2.75	223	6	.92	74.33	2.00
5587	7.0	7.00	1078	627	13	6.75	523	12	.96	74.71	1.71	2j5587	3	3.00	79	3	1.00	26.33	1.00	3	2.75	324	3	.92	108.00	1.00	1	1.00	120	6	1.00	120.00	6.00
5587	9.0	9.00	681	340	16	9.00	340	16	1.00	37.78	1.78	1v5587	3	3.00	66	0	1.00	22.00	.00	3	3.00	126	4	1.00	42.00	1.33	3	3.00	148	12	1.00	49.33	4.00
5602	9.0	9.00	908	568	9	9.00	568	9	1.00	63.11	1.00	2j5602	3	3.00	111	2	1.00	37.00	.67	3	3.00	283	4	1.00	94.33	1.33	3	3.00	174	3	1.00	58.00	1.00
5602	9.0	9.00	513	264	12	9.00	264	12	1.00	29.33	1.33	1v5602	3	3.00	51	0	1.00	17.00	.00	3	3.00	80	4	1.00	26.67	1.33	3	3.00	133	8	1.00	44.33	2.67
5607	8.0	8.75	881	511	7	8.00	506	7	1.00	63.25	.88	2j5607	3	3.00	110	2	1.00	36.67	.67	3	3.00	214	3	1.00	71.33	1.00	2	2.00	182	2	1.00	91.00	1.00
5607	8.0	9.00	634	291	9	8.00	268	9	1.00	33.50	1.13	1v5607	2	2.00	42	1	1.00	21.00	.50	3	3.00	111	3	1.00	37.00	1.00	3	3.00	115	5	1.00	38.33	1.67
5621	8.0	7.25	950	597	9	6.50	518	8	.81	64.75	1.00	1j5621	3	2.50	95	1	.83	31.67	.33	3	2.50	234	5	.83	78.00	1.67	2	1.50	189	2	.75	94.50	1.00
5621	9.0	8.00	496	224	12	8.00	224	12	.89	24.89	1.33	2v5621	3	2.75	46	0	.92	15.33	.00	3	3.00	61	3	1.00	20.33	1.00	3	2.25	117	9	.75	39.00	3.00
5632	8.0	8.75	1438	872	20	7.75	786	17	.97	98.25	2.13	1j5632	2	2.00	210	7	1.00	105.00	3.50	3	2.75	356	6	.92	118.67	2.00	3	3.00	220	4	1.00	73.33	1.33
5632	9.0	9.00	539	253	13	9.00	253	13	1.00	28.11	1.44	2v5632	3	3.00	55	0	1.00	18.33	.00	3	3.00	74	5	1.00	24.67	1.67	3	3.00	124	8	1.00	41.33	2.67
5642	9.0	8.75	1173	705	12	8.75	705	12	.97	78.33	1.33	1j5642	3	2.75	119	3	.92	39.67	1.00	3	3.00	221	3	1.00	73.67	1.00	3	3.00	365	6	1.00	121.67	2.00
5642	9.0	9.00	555	286	14	9.00	286	14	1.00	31.78	1.56	2v5642	3	3.00	64	1	1.00	21.33	.33	3	3.00	90	4	1.00	30.00	1.33	3	3.00	132	9	1.00	44.00	3.00
5649	9.0	9.00	996	597	7	9.00	597	7	1.00	66.33	.78	2j5649	3	3.00	146	1	1.00	48.67	.33	3	3.00	228	3	1.00	76.00	1.00	3	3.00	223	3	1.00	74.33	1.00
5649	9.0	8.75	655	368	11	8.75	368	11	.97	40.89	1.22	1v5649	3	3.00	52	0	1.00	17.33	.00	3	2.75	159	4	.92	53.00	1.33	3	3.00	157	7	1.00	52.33	2.33
5679	9.0	7.75	1106	643	12	7.75	643	12	.86	71.44	1.33	1j5679	3	2.25	129	1	.75	43.00	.33	3	3.00	290	7	1.00	96.67	2.33	3	2.50	224	4	.83	74.67	1.33
5679	8.0	8.75	722	352	20	7.75	271	11	.97	33.88	1.38	2v5679	3	2.75	77	1	.92	25.67	.33	3	3.00	141	6	1.00	47.00	2.00	2	2.00	53	4	1.00	26.50	2.00
5687	8.0	7.25	1102	620	14	6.75	543	13	.84	67.88	1.63	1j5687	3	2.00	210	6	.67	70.00	2.00	2	1.75	91	2	.88	45.50	1.00	3	3.00	242	5	1.00	80.67	1.67
5687	8.0	9.00	450	198	12	8.00	186	11	1.00	23.25	1.38	2v5687	3	3.00	53	0	1.00	17.67	.00	2	2.00	34	2	1.00	17.00	1.00	3	3.00	99	9	1.00	33.00	3.00
5698	8.0	8.25	1345	713	9	8.00	615	8	1.00	76.88	1.00	1j5698	3	3.00	126	1	1.00	42.00	.33	2	2.00	300	4	1.00	150.00	2.00	3	3.00	189	3	1.00	63.00	1.00
5698	9.0	8.75	592	280	16	8.75	280	16	.97	31.11	1.78	2v5698	3	2.75	89	2	.92	29.67	.67	3	3.00	74	3	1.00	24.67	1.00	3	3.00	117	11	1.00	39.00	3.67
5699	7.0	7.25	1664	1093	19	6.75	770	13	.96	110.00	1.86	1j5699	3	3.00	389	6	1.00	129.67	2.00	3	2.75	239	4	.92	79.67	1.33	1	1.00	142	3	1.00	142.00	3.00
5699	9.0	8.50	554	254	13	8.50	254	13	.94	28.22	1.44	2v5699	3	2.75	86	2	.92	28.67	.67	3	3.00	55	3	1.00	18.33	1.00	3	2.75	113	8	.92	37.67	2.67
5715	9.0	8.25	1437	951	14	8.25	951	14	.92	105.67	1.56	2j5715	3	2.50	328	3	.83	109.33	1.00	3	2.75	366	6	.92	122.00	2.00	3	3.00	257	5	1.00	85.67	1.67
5715	8.0	8.25	961	427	15	7.75	366	14	.97	45.75	1.75	1v5715	3	2.75	149	4	.92	49.67	1.33	2	2.00	61	1	1.00	30.50	.50	3	3.00	156	9	1.00	52.00	3.00
5736	9.0	6.50	1103	652	9	6.50	652	9	.72	72.44	1.00	1j5736	3	2.00	108	2	.67	36.00	.67	3	2.25	359	4	.75	119.67	1.33	3	2.25	185	3	.75	61.67	1.00
5736	8.0	8.50	518	255	11	7.50	221	10	.94	27.63	1.25	2v5736	2	1.75	18	0	.88	9.00	.00	3	3.00	61	3	1.00	20.33	1.00	3	2.75	142	7	.92	47.33	2.33
5737	7.0	9.00	1906	1179	12	7.00	934	10	1.00	133.43	1.43	1j5737	3	3.00	220	1	1.00	73.33	.33	3	3.00	482	5	1.00	160.67	1.67	1	1.00	232	4	1.00	232.00	4.00
5737	9.0	9.00	614	324	13	9.00	324	13	1.00	36.00	1.44	2v5737	3	3.00	69	0	1.00	23.00	.00	3	3.00	79	3	1.00	26.33	1.00	3	3.00	176	10	1.00	58.67	3.33
5754	6.0	7.00	1493	765	13	5.50	461	6	.92	76.83	1.00	2j5754	3	2.75	176	2	.92	58.67	.67	2	1.75	209	3	.88	104.50	1.50	1	1.00	76	1	1.00	76.00	1.00
5754	9.0	9.00	653	316	11	9.00	316	11	1.00	35.11	1.22	1v5754	3	3.00	88	1	1.00	29.33	.33	3	3.00	70	3	1.00	23.33	1.00	3	3.00	158	7	1.00	52.67	2.33
5778	7.0	7.75	1053	616	9	6.25	454	6	.89	64.86	.86	2j5778	3	3.00	157	1	1.00	52.33	.33	3	2.25	243	4	.75	81.00	1.33	1	1.00	54	1	1.00	54.00	1.00
5778	9.0	8.75	741	405	14	8.75	405	14	.97	45.00	1.56	1v5778	3	3.00	160	1	1.00	53.33	.33	3	2.75	65	3	.92	21.67	1.00	3	3.00	180	10	1.00	60.00	3.33
5801	9.0	9.00	1574	1020	15	9.00	1020	15	1.00	113.33	1.67	1j5801	3	3.00	242	3	1.00	80.67	1.00	3	3.00	288	4	1.00	96.00	1.33	3	3.00	490	8	1.00	163.33	2.67
5801	9.0	9.00	676	329	11	9.00	329	11	1.00	36.56	1.22	2v5801	3	3.00	97	1	1.00	32.33	.33	3	3.00	124	4	1.00	41.33	1.33	3	3.00	108	6	1.00	36.00	2.00
5810	5.0	5.50	1670	1086	16	3.50	567	8	.70	113.40	1.60	1j5810	1	1.00	107	2	1.00	107.00	2.00	2	2.00	206	2	1.00	103.00	1.00	2	.50	254	4	.25	127.00	2.00
5810	9.0	9.00	697	307	13	9.00	307	13	1.00	34.11	1.44	2v5810	3	3.00	47	0	1.00	15.67	.00	3	3.00	97	5	1.00	32.33	1.67	3	3.00	163	8	1.00	54.33	2.67
5812	9.0	9.00	1306	786	13	9.00	786	13	1.00	87.33	1.44	2j5812	3	3.00	136	1	1.00	45.33	.33	3	3.00	334	7	1.00	111.33	2.33	3	3.00	316	5	1.00	105.33	1.67
5812	9.0	9.00	663	417	7	9.00	417	7	1.00	46.33	.78	1v5812	3	3.00	206	1	1.00	68.67	.33	3	3.00	86	2	1.00	28.67	.67	3	3.00	125	4	1.00	41.67	1.33
5816	9.0	4.50	1417	836	12	4.50	836	12	.50	92.89	1.33	2j5816	3	2.50	91	1	.83	30.33	.33	3	.75	385	4	.25	128.33	1.33	3	1.25	360	7	.42	120.00	2.33
5816	9.0	7.00	661	343	14	7.00	343	14	.78	38.11	1.56	1v5816	3	1.50	138	1	.50	46.00	.33	3	3.00	74	3	1.00	24.67	1.00	3	2.50	131	10	.83	43.67	3.33
5836	5.0	4.75	1873	667	12	4.50	300	5	.90	60.00	1.00	1j5836	1	1.00	78	1	1.00	78.00	1.00	2	1.75	125	2	.88	62.50	1.00	2	1.75	97	2	.88	48.50	1.00
5836	5.0	4.50	566	310	14	3.25	166	7	.65	33.20	1.40	2v5836	1	.50	39	1	.50	39.00	1.00	1	.00	28	0	.00	28.00	.00	3	2.75	99	6	.92	33.00	2.00
5856	9.0	7.00	1257	805	16	7.00	805	16	.78	89.44	1.78	1j5856	3	2.75	255	7	.92	85.00	2.33	3	2.25	202	4	.75	67.33	1.33	3	2.00	348	5	.67	116.00	1.67
5856	9.0	8.75	513	259	15	8.75	259	15	.97	28.78	1.67	2v5856	3	3.00	81	2	1.00	27.00	.67	3	3.00	51	3	1.00	17.00	1.00	3	2.75	127	10	.92	42.33	3.33
5873	8.0	9.00	1315	680	13	8.00	595	11	1.00	74.38	1.38	2j5873	3	3.00	199	4	1.00	66.33	1.33	2	2.00	159	3	1.00	79.50	1.50	3	3.00	237	4	1.00	79.00	1.33
5873	8.0	9.00	641	340	13	8.00	249	7	1.00	31.13	.88	1v5873	3	3.00	101	2	1.00	33.67	.67	3	3.00	84	3	1.00	28.00	1.00	2	2.00	64	2	1.00	32.00	1.00
5875	9.0	6.00	1204	742	13	6.00	742	13	.67	82.44	1.44	2j5875	3	2.25	127	2	.75	42.33	.67	3	2.25	331	6	.75	110.33	2.00	3	1.50	284	5	.50	94.67	1.67
5875	9.0	6.50	1053	579	16	6.50	579	16	.72	64.33	1.78	1v5875	3	2.25	207	2	.75	69.00	.67	3	2.25	157	7	.75	52.33	2.33	3	2.00	215	7	.67	71.67	2.33
5886	5.0	6.25	1043	635	10	4.00	261	5	.80	52.20	1.00	1j5886	3	2.00	137	2	.67	45.67	.67	1	1.00	71	2	1.00	71.00	2.00	1	1.00	53	1	1.00	53.00	1.00
5886	9.0	8.75	422	198	9	8.75	198	9	.97	22.00	1.00	2v5886	3	3.00	41	0	1.00	13.67	.00	3	2.75	63	3	.92	21.00	1.00	3	3.00	94	6	1.00	31.33	2.00
5892	8.0	7.75	1384	838	7	6.75	734	6	.84	91.75	.75	1j5892	2	2.00	54	0	1.00	27.00	.00	3	2.75	270	3	.92	90.00	1.00	3	2.00	410	3	.67	136.67	1.00
5892	9.0	8.75	590	295	14	8.75	295	14	.97	32.78	1.56	2v5892	3	2.75	85	2	.92	28.33	.67	3	3.00	74	4	1.00	24.67	1.33	3	3.00	136	8	1.00	45.33	2.67
5893	6.0	2.50	1263	739	12	2.00	522	10	.33	87.00	1.67	1j5893	2	1.25	88	1	.63	44.00	.50	3	.75	351	7	.25	117.00	2.33	1	.00	83	2	.00	83.00	2.00
5893	8.0	4.25	667	385	8	4.00	345	8	.50	43.13	1.00	2v5893	3	1.25	52	0	.42	17.33	.00	3	1.75	123	3	.58	41.00	1.00	2	1.00	170	5	.50	85.00	2.50
5898	9.0	9.00	1017	465	11	9.00	465	11	1.00	51.67	1.22	1j5898	3	3.00	134	4	1.00	44.67	1.33	3	3.00	124	3	1.00	41.33	1.00	3	3.00	207	4	1.00	69.00	1.33
5898	9.0	8.50	459	186	11	8.50	186	11	.94	20.67	1.22	2v5898	3	3.00	38	0	1.00	12.67	.00	3	3.00	47	3	1.00	15.67	1.00	3	2.50	101	8	.83	33.67	2.67
5902	5.0	7.00	1255	713	25	4.75	352	16	.95	70.40	3.20	2j5902	2	1.75	171	8	.88	85.50	4.00	2	2.00	79	4	1.00	39.50	2.00	1	1.00	102	4	1.00	102.00	4.00
5902	8.0	9.00	690	316	13	8.00	243	13	1.00	30.38	1.63	1v5902	2	2.00	43	1	1.00	21.50	.50	3	3.00	78	5	1.00	26.00	1.67	3	3.00	122	7	1.00	40.67	2.33

Table 38 Subject Totals and Correct Task Averages

SUBJ	ORDT	EAS	HLP	FST	INF	GUI	LRN	FUN	USE	SRV
2201	1JIM	.00	.25	.50	.50	.25	.25	.00	.50	2.25
2201	2VIT	.75	.75	.50	.75	.75	.75	.75	.75	5.75
2598	2JIM	.25	.75	.00	.50	.50	.50	.00	.75	3.25
2598	1VIT	1.00	.75	1.00	.50	.75	.50	.75	.75	6.00
2961	1JIM	.25	.50	.25	.75	.50	.25	.00	.50	3.00
2961	2VIT	1.00	.75	1.00	1.00	.75	.75	.75	.75	6.75
3749	2JIM	.75	1.00	1.00	1.00	.50	1.00	.50	1.00	6.75
3749	1VIT	1.00	.75	.75	.75	1.00	.75	.75	.75	6.50
3779	2JIM	.00	.00	.00	.00	.00	.00	.00	.00	.00
3779	1VIT	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	8.00
3793	1JIM	.25	.25	.00	.75	.25	.25	.25	.25	2.25
3793	2VIT	1.00	.75	1.00	.75	1.00	.75	.75	1.00	7.00
3795	1JIM	.25	.25	.00	.00	.25	.25	.00	.25	1.25
3795	2VIT	1.00	.75	1.00	1.00	.75	.75	.50	.75	6.50
5161	2JIM	.25	.50	.25	.50	.50	.50	.25	.50	3.25
5161	1VIT	.00	.75	.75	.75	1.00	.75	.50	.75	5.25
5407	2JIM	.00	.25	.25	.75	.50	.50	.00	.50	2.75
5407	1VIT	.75	.50	.75	.75	.50	.75	.50	.75	5.25
5424	1JIM	.50	.75	.25	.50	.50	.75	.50	.75	4.50
5424	2VIT	.75	.75	.50	.50	.75	.75	.75	.75	5.50
5516	1JIM	.50	.75	.25	.50	.25	.75	.50	.75	4.25
5516	2VIT	.75	.75	1.00	.75	.75	.75	.75	.75	6.25
5546	2JIM	.75	.75	.25	1.00	.25	.75	.50	.75	5.00
5546	1VIT	.75	.75	.75	.75	.75	.50	.50	.75	5.50
5582	2JIM	.50	.75	.75	.75	.75	.75	.75	.75	5.75
5582	1VIT	1.00	.75	.25	.75	.75	.50	.50	.75	5.25
5587	2JIM	.00	.25	.00	.25	.25	.50	.00	.50	1.75
5587	1VIT	1.00	1.00	1.00	1.00	1.00	1.00	.75	1.00	7.75
5602	2JIM	.00	.00	.00	.00	.00	.00	.00	.00	.00
5602	1VIT	.00	.00	.25	.50	.00	.00	.25	.00	1.00
5607	2JIM	.75	1.00	.75	.75	.75	.75	.50	.75	6.00
5607	1VIT	1.00	.75	.75	.75	1.00	1.00	.75	1.00	7.00
5621	1JIM	.25	.75	.25	.75	.50	.25	.25	.75	3.75
5621	2VIT	.75	1.00	.75	.25	.50	.75	1.00	.75	5.75
5632	1JIM	.25	.50	.25	.75	.25	.50	.00	.25	2.75
5632	2VIT	.25	.75	.75	.50	.75	.75	.75	.75	5.25
5642	1JIM	.00	.75	.00	1.00	.50	.75	.50	1.00	4.50
5642	2VIT	1.00	.75	1.00	1.00	1.00	1.00	.50	1.00	7.25
5649	2JIM	.75	.75	.25	1.00	.75	.50	.50	.50	5.00
5649	1VIT	.75	.00	1.00	1.00	.75	.75	1.00	.50	5.75
5679	1JIM	1.00	1.00	.75	.75	.50	.75	.75	.50	6.00
5679	2VIT	.75	.75	.25	.75	.50	.75	.25	.50	4.50
5687	1JIM	.50	.25	.25	.25	.25	.25	.00	.50	2.25
5687	2VIT	.75	.75	1.00	1.00	.75	.75	1.00	.75	6.75
5698	1JIM	.50	.75	.75	.75	.25	.75	.25	.50	4.50
5698	2VIT	.75	.75	.25	.75	1.00	.75	.75	.75	5.75
5699	1JIM	.25	.50	.25	.25	.50	.75	.50	.75	3.75
5699	2VIT	1.00	.75	1.00	.75	.75	.75	.75	.75	6.50
5715	2JIM	.00	.25	.00	.75	.50	.25	.00	.25	2.00
5715	1VIT	.75	.50	.75	.75	.50	.50	.50	.75	5.00
5736	1JIM	.25	.75	.25	.75	.25	.25	.25	.75	3.50
5736	2VIT	.25	.75	.75	.25	.50	.75	.75	.75	4.75
5737	1JIM	.00	.25	.00	1.00	.75	.75	.25	.75	3.75
5737	2VIT	.75	.75	1.00	1.00	.75	.75	.25	.75	6.00
5754	2JIM	.00	.25	.00	.00	.25	.75	.00	1.00	2.25
5754	1VIT	1.00	.75	1.00	1.00	.75	1.00	1.00	1.00	7.50
5778	2JIM	.75	.75	.50	.75	.50	.75	.75	.75	5.50
5778	1VIT	.75	.75	.50	.75	.50	.75	.25	.75	5.00
5801	1JIM	.50	.25	.25	.50	.25	.50	.25	.75	3.25
5801	2VIT	1.00	.75	1.00	.50	.75	.50	.75	.75	6.00
5810	1JIM	.25	.50	.00	.00	.00	.25	.00	.00	1.00
5810	2VIT	1.00	.75	1.00	1.00	1.00	1.00	1.00	1.00	7.75
5812	2JIM	.75	.75	.25	.75	.50	.50	.25	.75	4.50
5812	1VIT	.75	.75	.75	.75	.75	.50	.75	.75	5.75
5816	2JIM	.00	.25	.00	.50	.50	.25	.00	.50	2.00
5816	1VIT	.75	.75	1.00	.50	.50	.75	1.00	.50	5.75
5836	1JIM	.00	.00	.75	.75	.00	.00	.00	.00	1.50
5836	2VIT	.00	.00	.50	.00	.50	.00	.00	.00	1.00
5856	1JIM	.75	.75	1.00	.75	.25	.50	.50	.75	5.25
5856	2VIT	.25	.75	.25	.75	.75	.75	.50	.75	4.75
5873	2JIM	.75	.50	1.00	.75	.75	.50	.75	.75	5.75
5873	1VIT	.75	.50	.25	.75	.50	.50	.50	.50	4.25
5875	2JIM	.25	.50	.75	.25	.50	.25	.00	.25	2.75
5875	1VIT	.75	.75	.75	.75	.75	.75	.75	.75	6.00
5886	1JIM	.25	.50	.00	.25	.00	.00	.25	.25	1.50
5886	2VIT	1.00	.75	1.00	.75	1.00	1.00	.75	.75	7.00
5892	1JIM	.75	.75	.50	.75	.25	.50	.25	.75	4.50
5892	2VIT	1.00	1.00	1.00	1.00	1.00	.75	1.00	.75	7.50
5893	1JIM	.00	.25	.00	.25	.25	.25	.00	.00	1.00
5893	2VIT	.25	.50	.75	.50	.50	.75	.50	.50	4.25
5898	1JIM	.75	.75	1.00	.75	.25	.75	.75	.75	5.75
5898	2VIT	.75	.00	.00	.75	.75	.50	.00	.25	3.00
5902	2JIM	.25	.50	.00	.50	.50	.75	.00	.50	3.00
5902	1VIT	1.00	1.00	1.00	1.00	.75	.75	.75	.75	7.00

Table 39 Subject Satisfaction Survey Responses

APPENDIX K: QUALITATIVE DATA

This appendix contains the raw comments from the subjects of this research, Question # 17 requested feedback about VITAMIN, #18 about JIMI, and #19 overall comments.

SUBJ	Q#	Comments
2201	17	THIS INTERFACE WAS MUCH EASIER TO USE. I WAS ABLE TO VISUALIZE HOW THE DATA REALATED. DEFINITLEY THE BETTER OF THE TWO INTERFACES.
2201	18	THIS INTERFACE WAS DIFFICULT AND CONFUSING. IT WAS HARD TO UNDERTAND AND SEE THE RELATIONSHIP OF THE DATA.
2201	19	VITAMIN I FOUND TO BE THE BETTER OF THE TWO INTERFACES.
2598	17	The vitamin system was easy to use. The data was easily reached and it came quickly. The red and green plusses were more stimulating than numbers and letters.
2598	18	The JIMI system was very hard to use. the information was hard to find and errors were made more frequently. the information was easily read, but it was hard to find. the information was also slow to come up.
2598	19	The vitamin system should be used instead of the jimi because it is more user-friendly and the information is faster to arrive. the vitamin also makes for less keystroke errors.
2961	17	The VITAMIN System GUI was the easier of the two programs to use. I liked the fact that the information was all laid out before you and if you wanted to access another category then you just clicked on the plus or minus button. You could also see all the information of the ships laid out in a more informative format. I found this program to be easy to use and not hard at all to learn.
2961	18	The JIMI program was hard at first to understand. It was difficult to understand exactly what the question provided was asking. After some practice the program wasn't that difficult. I think it would be easy to make a mistake reading data in this program. You also had to jump back and forth between screens, which could lead to confusion.
2961	19	Overall, I like the VITAMIN program better of the two. It was clear and easy to use. The information was laid out in a way that was easy to interpret. The plus and minus buttons also made the program more easy to use.
3749	17	I found that the process in finding out the information in this system was quite simple. However, it was difficult, painstaking, and time-consuming to compare the different companies and ships to one another. The tutorial to VITAMIN made it extremely simple to use and ifnd out answers, however once those answers were drawn up it was a little bit more of a difficult process, causing me to constantly scroll up and scroll down.
3749	18	JIMI was at first a little more difficult and cumbersome to use than was VITAMIN. This was only because it was not as simple as the plus sign system and I actually had to type in data. However, after the first few questions I felt very confident in how to use the system and I enjoyed it greatly. The main thing that I liked about JIMI was that it was very simple to compare the data of the different ships, companies and other components once they were drawn up.
3749	19	Both of these systems are accurate it seems and quite effective. I did not find either of them to be difficult to use, and with the proper tutorial and training and individual could use either program. For the sake of simplicity and lack of complication VITAMIN would be more effective, however with a more experienced individual looking for expediency of data JIMI would seem to me to be the ideal program.

SUBJ	Q#	Comments
3779	17	It is very user friendly. If you make a mistake, it is easy to see where you went wrong, so you can go back and correct it. There is no problem with typing because it is all done through the use of a mouse. Spelling and format errors are not a problem.
3779	18	I personally hate the JIMI way of doing things. It was a giant pain in the butt because you had to make sure everything was grammatically correct in order to get a query. It was more confusing on what kind of query to make. It took up more of my time than I would have liked it to. I wasn't very user friendly unless you knew the system very well or could easily adapt. I don't like it. I, like many other Americans, am in a hurry.
3779	19	If the JIMI methods is chosen to be used over the VITAMIN method, then I quit.
3793	17	It was much easier to use than the JIMI. The only bad thing is that on the JIMI you can display all of values of the ships easier. On the VITAMIN you have to go through and look at all of the ships to compare them. In the JIMI it does it for you.
3793	18	It took me a while to get use to. In the begining it was hard to understand what the questions were asking and how to get the answer. One good thing is that the JIMI displayed all of the vaules for the ships unlike the VITAMIN in which you had to look up all of the ships separately
3793	19	I think the VITAMIN is eaier to use and I would recommend using that one over the JIMI.
3795	17	The VITAMIN sytem was much more user friendly. I moved through the questions at a much faster rate with the VITAMIN system. This also may have been due to my adapting to the data, but I still feel that the VITAMIN sytem is more efficient overall. I was, however, a little surprised at the slow speed of the VITAMIN system when dealing with just numerical data.
3795	18	I became a little frustrated with the JIMI system almost immediately. For someone you is not used to the language utilized, the system takes a little while to learn. Once I got the hang of the system, I was able to manuever a little quicker, but I feel that the VITAMIN system suceeds in overcoming these initial misunderstandings where the JIMI system does not.
3795	19	The whole process was very foreign for me at first. I did not initially understand the instructions or the purpose of the two systems. By the end of the questions, however, I feel that I have gained a basic understanding of the data. The VITAMIN system and assistance from the facilitator had more of an impact on this than anything else.
5161	17	Much easyer to use, organize, and tell what is going on enough to answer the questions without having a real good computer understanding and background. Vitamin would be much better for new workers because it is very easy to learn and understand.
5161	18	JIMI requires time to think about what is what varible and their meanings. It also takes too much time to input the funtions to run the queries. It only takes one space typed inside the parenthesis to throw off the queries and get bogus or so data back.
5161	19	Obviously the Vitamin system is much easyer to use and understand.
5407	17	
5407	18	I never really was able to understand how to use it. However, I am sure that once I could understand it, it would become very useful to me.
5407	19	
5424	17	I felt that the Vitamin program was far easier to use than the JIMI program, Both systems are easy to use, but it takes someone a little longer to figure out excatally what the JIMI program wants, and how to get. Where as the Vitamin program was all click and choose, easier to use and easier to catch on to.
5424	18	look on the page before
5424	19	
5516	17	The VITAMIN system was much easier to use and see what you were looking at.

SUBJ	Q#	Comments
5516	18	It was too ambiguous. You couldn't follow the data as easy as you could on the VITAMIN system.
5516	19	I found both systems useful in finding data on supplies. However, the VITAMIN system was quicker and more affective.
5546	17	Once learning how to work the actual program and being able to use it to answer the specific questions, teh VITAMIN system was pretty easy to use.
5546	18	JIMI had grouped the different companies and ships to where it was easy to get whatever info. was needed.
5546	19	
5582	17	VITAMIN was very easy to use. The only problem is that some questions required either all of the plusses to be opened or that they be opened and closed one at a time. The first case made it difficult to easliy view and comprehend the data, and the second case was just a pain in the neck.
5582	18	JIMI required quite a bit more concentration than VITAMIN. Personally, it took longer to use JIMI than VITAMIN. However, I believe that with time and practise, JIMI would be a much more efficient and speedy tool. I also thought that JIMI was better because it allowed the operator to go directly to the data needed. I did not have to open multiple plusses.
5582	19	JIMI was more difficult to operate, initially, but with practice I believe it is the most efficient and comprehensible of the two.
5587	17	The VITAMIN System GUI was a much better system than the JIMI system for several reasons. One being that the VITAMIN system had everything layed out so it was easy to see where you wanted to go .
5587	18	The JIMI system was very confusing and sometimes a hassle to work. It was sometime difficult to find the data needed. This was probably because it was not layed out clearly. The outline form of the VITAMIN system was much more clear and easy to use.
5587	19	If I were to choose between the two systems, I would definitely choose the VITAMIN system because it would eliminate several common errors. When I was scrolling through the two systems, I noticed that it would be very easy to make a data mistake while using the JIMI system. This would probably cause some major problems.
5602	17	IT WAS A LITTLE SLOW AND THE INFORMATION IS ALL JUMBLED UP. THE PLUSSES AND MINUSES MAKE FOR EAQST ACCESS BUT ASS MOREOF THE PLUSSES AND MINUSES ARE OPENED THE MORE CUNFUSING IT GETS. oN THE OTHER HAND THE INFORMATION IS VERY ORGANIZED, BUT MAYBE INTEGRATING MODULES INTO THE PROGRAM WOULD ENHANCE ITS CAPABILITIES.
5602	18	VERY ORGANIZED, THE DATA IS VERY EASILY ACCESSIBLE AND UNDERSTANDABLE. YOU GET EXACTLY WHAT YOU ASK FOR IN ONE EASY STEP. I CAN'T SAY VERY MUCH ABOUT THE JAVA PROGRAM BEACUSE IT IS TOTALLY FLUID. IT WORKS VERY WELL AND CAN BE COMPATIBLE TO ANY SITUATION
5602	19	BOTH PROGRAMS WERE GOOD, ALTHOUGH THEY DIFFERED IN SOME AREAS, THEY BOTH CAN BE INTEGRATED AND USED VERY PRODUCTIVELY. NIETHER ON EIS BETTYER THAN THE OTHER, THEY JUST PRESENT THE DATA A LITTLE DIFFERENT.
5607	17	The VITAMIN System Gui was extremly useful and extremely easy to use. It was a very nice System in that it provided a visual classification that was easy to comprehend.
5607	18	The JIMI was slightly less usable as the VITAMIN System. It did not have the visual impact that the VITAMIN system had. It was also more time consuming to determine information as it often required manual typing of information.
5607	19	The VITAMIN system was the superior system. It was more usable and provided more effecient information. It was easier to comprehend because of its visual nature.

SUBJ	Q#	Comments
5621	17	The VITAMIN System seemed easier to use. It was faster because there was less thought that had to take place as to where I needed to go to get the information needed. That seemed very important to me because I was more sure that the information that I had retrieved was right. However, I feel as though I could get more information off of the JIMI system.
5621	18	This was more difficult to use and more tedious. I first needed to find out which place I had to go to retrieve the correct information, and then I needed to be sure to put the words in the correct place. Once I was able to do that, the system worked faster. The problem is that it did take me a little longer to input the words (especially since it was easier to make a mistake). It was not as user friendly.
5621	19	The colored buttons on the VITAMIN system helped also because a soldier would be able to determine approximately the percent for each place just by the color of the plus or minus sign. I felt the VITAMIN system was easier to use and faster. However, I believe I could have retreived more information off of the JIMI system.
5632	17	This system allowed me to find the information faster than the JIMI system when asking general information, but the JIMI system was faster when finding specific information.
5632	18	This is hard to pick up and learn versus the VITAMIN system which was very easy to learn. This system is good for finding specific information.
5632	19	I think the VITAMIN system is overall better, for the questions asked I was able to find the answer far faster using VITAMIN versus JIMI. VITAMIN is not perfect though. It takes a while to find specific information on the VITAMIN system, a bigger viewer interface could allow the user to view more information and not have to open and close as much.
5642	17	It was much easier to use thn the JIMI. It was also much easier to learn and allowed to find the data much more quickly. I am not sure though if the JIMI system would be faster if I had a longer time to learn how to use the queries and get accustomed to the terminology.
5642	18	The JIMI system is as effecient at finding the data as the VITAMIN, but it took longer to learn and apply. At first it was very confusing. One thing I can say about this system is that my perception of how quickly and easily it can be run is not accurate because the JIMI system was practice before the VITAMIN system. Therefore, my opinion is quite mixed.
5642	19	It is hard to determine if one system is easier to learn than the others or that i mattered which system I learned first.
5649	17	THIS SYSTEM WAS A LOT MORE USER FRIENDLY AND ALLOWES AN INDIVIDUAL TO WORK FASTER WITHOUT HAVE TO WAIT ON THE COMPUTER TO RELOAD OR QUERY FOR EVERY SELECTION.
5649	18	IT CAUSES A PERSON TO RECALL MORE SIGNS OR ABBREVIATION, BUT IT CAN BE EFFECTIVE SINCE IT WILL GO DIRECTLY TO THE INFORMATION ONE IS LOOKING FOR, BUT TYPING TAKES TIME AND CALL LEAD TO ERRORS WHICH CAN THROW OFF YOUR INFORMATION BEING SOUGHT OUT.
5649	19	THE VITAMIN SYSTEM SHOULD BE IMPLEMENTED AFTER EVERYONE HAS LEARNED TO USE THE JIMI SYSTEM FIRST SO ONE CAN HAVE A BETTER UNDERSTANDING OF THE MATERIAL AND HOW TO LOOK IT UP.
5679	17	The VITAMIN System GUI was a little more complex and did was not very clear. I did not like having to type in different names and titles to find out information.
5679	18	The JIMI was easier to use and was self explanatory. The interface was more fluent and easier to read and at a level that was not hard to comprehend.
5679	19	The JIMI seemed to be quicker because of the arrows that one would click on instead of having to type in new titles.
5687	17	I thought the vitamin format was very easy to use. Some how it seemed like a computer game where you are searching for something and you have to go into doors to find. I found it very enjoyable to use.

SUBJ	Q#	Comments
5687	18	I disliked the JIMI format. I was very cumbersome and annoying to have to type in different answers to search for every time. It was also difficult to understand what some of the symbols stood for.
5687	19	The VITAMIN format was much better that the JIMI format. I found VITAMIN to be enjoyable to work with and never got frustrated or felt my answers were incorrect. On the other hand the JIMI format was very difficult to understand and I got frustrated with it many times. Also I was unsure about many of my answers in the JIMI format.
5698	17	The Vitamin system was easier to use. I figured out how to use it in a brief amount of time. It was easier to click on your choices rather than type them in.
5698	18	JIMI was difficult to learn how to use.
5698	19	I felt that VITAMIN was a better program to use.
5699	17	The VITAMIN system was by far easier to use than the JIMI. One click of the mouse and you had results that were labeled and in easy to read columns.
5699	18	Difficult, took a while to figure out where and how to enter the information. Results were also difficult to analyze. Also difficult because it is case sensitive. Most likely have to be experienced (somewhat) to run the JIMI system.
5699	19	Good evaluation to choose which type of programs are more effective and user friendly.
5715	17	VITAMIN was a lot easier to use than JIMI. Once I understood what it was asking the data was easy and accesible. The program was a little slow.
5715	18	JIMI was very difficult to use. It took up to much time changing and typing. It was difficult to understand what I had to change at first.
5715	19	I believe that JIMI is not very useful while VITAMIN is. With a faster program and bigger box, information would be very easy to find using VITAMIN.
5736	17	The only time I thought this system was inferior was when I was asked to compare the percent fill of a specific unit. Other than that, it was a lot easier to use.
5736	18	Very difficult to figure out how this system worked initially.
5736	19	The VITAMIN system was clkoser to an actual computer system and was therefore a lot easier to undertsnad right off the bat, like menus on a computer. The JIMI system was usueful when I was looking for specific information, like the search feature of a Web Browser.
5737	17	VITAMIN is a much easier program to use because it is a single click for the information.
5737	18	The JIMI is just to complicated to use because you have to type everything in yourself instead of a single click like the VITAMIN.
5737	19	
5754	17	The VITAMIN System GUI was very simple to use. It was fast and did not require you to spend more time than required to look for any information.
5754	18	The JIMI was very complicated to use. For someone like myself to use the JIMI it is very frustrating. It is also time consuming because it required you to type in the information you were looking for.
5754	19	The VITAMIN System overall out performed the JIMI. For soldiers on the battle field, it is more convinient for them to push a button and get the information they need rather than sitting there trying to type in information which leaves room for errors.
5778	17	I THOUGHT BOTH PROCESSES WERE SLOW, BUT THE JIMI WAS MORE INTERACTIVE AND MORE ENJOYABLE TO USE. IT TOOK THE VITAMIN AGES TO PRODUCE A SEEMINGLY SMALL AMOUNT OF INFORMATION.
5778	18	THIS TOOL WAS MORE INTERACTIVE THAN THE VITAMIN BECAUSE OF ITS REQUIREMENT OF THE USER TO TYPE HIS/HER INTENT INTO THE PROGRAM. I LIKED THIS TOOL MORE.
5778	19	AS STATED ABOVE.
5801	17	The system allowed me to do the work faster and I comprehended the material easier because it was in a logical chart format that I am familiar with.

SUBJ	Q#	Comments
5801	18	After I learned how to use the system, it still slowed down the results of answering questions.
5801	19	none
5810	17	The VITAMIN System was much easier to use than the other system.
5810	18	Very confusing.
5810	19	Thye were both dificult to use until I got the hang of them at the end.
5812	17	The JIMI was a little more complicated to use because it required knowing under which sub-heading to enter the data and required altaring the information in the Query box. Both provided the same amount of information.
5812	18	
5812	19	
5816	17	Good
5816	18	Bad, very difficult to use and a very time consuming process
5816	19	I liked Vitamin a lot better
5836	17	It seemed too complex for my level of computer education. I'm sure if someone was trained to use the VITAMIN system, it would be fine. I had a hard time with it.
5836	18	Like VITAMIN, it seemed unnecessarily complicated.
5836	19	Both systems are to confusing.
5856	17	While the JIMI system worked faster the VITAMIN system was easier to figure out how to use and was easier to understand the data. I think that the VITAMIN system is better for someone who doesn't know how to use the data but since the JIMI was faster and gave all of the data up front it would be the better system to use after figuring it out.
5856	18	Again, I think the JIMI system would be the better system after knowing what the data meant.
5856	19	The JIMI was more useful.
5873	17	It was much easier to use, but it was also very slow and items were harder to find.
5873	18	This interface was much faster and accurate when it came to the information. At first, it was a little difficult to use, but once I understood exactly how to use it appropriately, it was very easy and helpful.
5873	19	I think that both are good interfaces, the VITAMIN interface would be better if it were faster and the JIMI is great, but difficult to understand at first.
5875	17	Liked the fact that all of the information was displayed on the screen. Clicking on the plus icons made it easy to access more data.
5875	18	I found it confusing to use. It was more compliocated than the VITAMIN program, and took me longer to use. I didn't like the fact that not all of the information was visible at one time. This made it difficult to compare data quickly.
5875	19	I found the VITAMIN program much more easier and effective.
5886	17	THE VITAMIN SYSTEM WAS MUCH EASIER TO MANIPULATE TO FIND THE REQUESTED INFORMATION. I FOUND IT SOMEWHAT DIFFICULT TO LINE UP THE RESULTS WITH THE CORRESPONDING COLUMNS AT THE TOP AS I GOT DEEPER AND DEEPER INTO A SHIPS INFORMATION. THIS WAS THE ONLY FLAW IN THE VITAMIN SYSTEM THAT I SAW.
5886	18	JIMI DID NOT PRESENT THE RESULTS OF MY QUERIES IN A READABLE MANNER WHATSOEVER. AS YOU CAN SEE BY THE CERTAINTY OF MY RESPONSES, I WAS NEVER REALLY SURE THAT I WAS READING THE CORRECT NUMBER FROM THE RESULTS. IT WAS DIFFICULT TO INTERPRET THEM.
5886	19	VITAMIN IS THE SUPERIOR SYSTEM. IT SHOULD PERHAPS BE PRESENTED IN A LARGER SIZED SQUARE WITH LINES SEPARATING THE COLUMNS TO MAKE IT EASIER TO READ. OTHER THAN THAT, THERE IS NO QUESTION THAT VITAMIN WAS EASY TO MANIPULATE AND INTERPRET.

SUBJ	Q#	Comments
5892	17	Excellent interface. Plus and minus buttons made it much faster and easier to use. Once speed of the computer is increased along with data retrieval rates the system will be very useful.
5892	18	While once I got t he hang of it it became easy to use there is too much room for error when typing. I believe that people would prefer to press a button (option) rather than type in their selection.
5892	19	Good experiment. Initially though, I would provide more instruction. I hope I helped in determining which system will be used in the future.
5893	17	This system was much more effective and helpful than the JIMI system. Much less confusing, and more logical.
5893	18	Made the whole process much more difficult than it needed to be. Very ineffective.
5893	19	The Vitamin system seems to be the much more effective of the two systems experimented with today
5898	17	VITAMIN has a GUI that is easy to understand at first but merely ends up being cumbersome in the end. I quickly learned how to use this interface yet it only seemed to slow me down in the later problems. It gives ease of use at the expense of speed. To me this was a big problem, I wanted the data fast and didn't care to wait around while the GUI expanded itself.
5898	18	JIMI's interface was much tougher to learn than VITAMIN's but once I figured it out and how to use it, I was able to quickly access the specific data I needed. It has a higher learning curve and was therefore harder in the beginning but once I was familiar with it I was able to use it very efficiently. I much prefer this quicker method than VITAMIN's interface.
5898	19	Although JIMI took a minute to learn it ended up being much faster and the minimal time to learn it was well invested. VITAMIN's interface proved only to slow me down as the problems became more complicated. JIMI allowed me to directly access the information I needed and not have to wait for each directory to open.

Table 40 Raw Qualitative Comments

APPENDIX L: STATISTICS

This appendix includes tables generated in SPSS for the t-tests. These tables include descriptive statistics and t-test results for the following: all questions, all questions washout, satisfaction survey, by type, by type washout, by question, and by question washout. Additionally, Box-plots are included for the following categories that are not reported in Chapter 4: by type washout, by question, and by question washout.

The independent samples t-test is used to test the hypotheses because there is no information about the population. For each hypothesis test, Levene's test of equality of variance (L. sig) is the first value analyzed. If this value is less than 0.05 then the "equality of variance assumption" is not met and separate variance test is used. Then the result is reported in the standard from, i.e. (t=-3.0, df=64, p<=.002) for the first test in Table 40. The t-statistic and the degrees of freedom are used to find the p-value. By comparing the specified α of 0.05 to the p-value (t_α) the null hypothesis is rejected since $t_\alpha < \alpha$. The p-value of the test is the smallest value of alpha for which the null hypothesis would be rejected. Note that the one-tailed p-value is used since the hypotheses have direction. The one-tailed p-value is calculated by dividing the 2-tailed value by two. Additional information presented in the tables include the differences fro the means and standard errors and the confidence intervals of the difference. The confidence interval means that the result is 95% sure to be no less than –1.227 and no greater than -.249, again using the first result from Table 40.

L.1 All Questions

L.1.1 Descriptive Statistics

Group Statistics

	Treatment	N	Mean	Std. Deviation	Std. Error Mean
AnswersAll	JIM	42	7.857	1.389	.214
	VIT	42	8.595	.767	.118
ConfCorMean	JIM	42	.8680	.1816	2.802E-02
	VIT	42	.9378	.1176	1.814E-02
UserTCorMean	JIM	42	82.2947	20.8906	3.2235
	VIT	42	34.4451	8.2271	1.2695
Usability Survey Sum	JIM	42	3.4107	1.7213	.2656
	VIT	42	5.7381	1.5211	.2347

Table 41 Descriptive Statistics – All Questions

L.1.2 Hypothesis Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AnswersAll	Equal variances assumed	13.525	.000	-3.014	82	.003	-.738	.245	-1.225	-.251
	Equal variances not assumed			-3.014	63.866	.004	-.738	.245	-1.227	-.249
ConfCorMean	Equal variances assumed	5.989	.017	-2.091	82	.040	-6.980E-02	3.338E-02	-.1362	-3.40E-03
	Equal variances not assumed			-2.091	70.248	.040	-6.980E-02	3.338E-02	-.1364	-3.23E-03
UserTCorMean	Equal variances assumed	33.128	.000	13.812	82	.000	47.8496	3.4645	40.9577	54.7415
	Equal variances not assumed			13.812	53.419	.000	47.8496	3.4645	40.9020	54.7971
Usability Survey Sum	Equal variances assumed	2.807	.098	-6.566	82	.000	-2.3274	.3545	-3.0325	-1.6223
	Equal variances not assumed			-6.566	80.778	.000	-2.3274	.3545	-3.0327	-1.6221

Table 42 T-Test – All Questions

L.2 All Questions Washout Assumption Test

The most conservative test of the washout assumption is to discard all but the first treatments and examine the degree to which the treatment effect is still valid. This analysis is reported in this section for each hypothesis.

L.2.1 Descriptive Statistics

Group Statistics

	Treatment	N	Mean	Std. Deviation	Std. Error Mean
AnswersAll	JIM	23	7.739	1.421	.296
	VIT	19	8.684	.478	.110
ConfCorMean	JIM	23	.8465	.1794	3.741E-02
	VIT	19	.9441	.1114	2.556E-02
UserTCorMean	JIM	23	86.1709	22.6438	4.7216
	VIT	19	38.7434	8.0927	1.8566
Usability Survey Sum	JIM	23	3.3043	1.4980	.3124
	VIT	19	5.7632	1.5399	.3533

Table 43 Descriptive Statistics – All Questions - Washout

L.2.2 Hypothesis Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AnswersAll	Equal variances assumed	14.304	.001	-2.767	40	.009	-.945	.342	-1.635	-.255
	Equal variances not assumed			-2.991	27.791	.006	-.945	.316	-1.592	-.298
ConfCorMean	Equal variances assumed	2.819	.101	-2.063	40	.046	-9.762E-02	4.731E-02	-.1933	-2.00E-03
	Equal variances not assumed			-2.155	37.376	.038	-9.762E-02	4.531E-02	-.1894	-5.85E-03
UserTCorMean	Equal variances assumed	18.344	.000	8.668	40	.000	47.4275	5.4714	36.3694	58.4856
	Equal variances not assumed			9.348	28.497	.000	47.4275	5.0735	37.0431	57.8119
Usability Survey Sum	Equal variances assumed	.613	.438	-5.228	40	.000	-2.4588	.4703	-3.4093	-1.5083
	Equal variances not assumed			-5.214	38.095	.000	-2.4588	.4716	-3.4134	-1.5042

Table 44 T-Test – All Questions - Washout

L.3 Satisfaction Survey

L.3.1 Descriptive Statistics

Group Statistics

	Treatment	N	Mean	Std. Deviation	Std. Error Mean
Ease of Use	JIM	42	.3512	.3027	4.671E-02
	VIT	42	.7500	.3024	4.667E-02
Helpful	JIM	42	.5179	.2782	4.293E-02
	VIT	42	.6845	.2534	3.910E-02
Quickly	JIM	42	.3214	.3372	5.202E-02
	VIT	42	.7500	.2922	4.508E-02
Information Quality	JIM	42	.5714	.3039	4.689E-02
	VIT	42	.7381	.2340	3.610E-02
Interface Quality	JIM	42	.3750	.2156	3.327E-02
	VIT	42	.7381	.2135	3.295E-02
Learn About Data	JIM	42	.4762	.2637	4.069E-02
	VIT	42	.7083	.2203	3.399E-02
Enjoyable	JIM	42	.2560	.2676	4.129E-02
	VIT	42	.6548	.2645	4.082E-02
Usable	JIM	42	.5417	.2865	4.420E-02
	VIT	42	.7143	.2248	3.468E-02
Usability Survey Sum	JIM	42	3.4107	1.7213	.2656
	VIT	42	5.7381	1.5211	.2347

Table 45 Descriptive Statistics – Satisfaction Survey

L.3.2 Hypothesis Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Ease of Use	Equal variances assumed	2.224	.140	-6.040	82	.000	-.3988	6.603E-02	-.5302	-.2675
	Equal variances not assumed			-6.040	82.000	.000	-.3988	6.603E-02	-.5302	-.2675
Helpful	Equal variances assumed	3.517	.064	-2.870	82	.005	-.1667	5.807E-02	-.2822	-5.11E-02
	Equal variances not assumed			-2.870	81.293	.005	-.1667	5.807E-02	-.2822	-5.11E-02
Quickly	Equal variances assumed	1.610	.208	-6.226	82	.000	-.4286	6.884E-02	-.5655	-.2916
	Equal variances not assumed			-6.226	80.374	.000	-.4286	6.884E-02	-.5656	-.2916
Information Quality	Equal variances assumed	6.907	.010	-2.817	82	.006	-.1667	5.917E-02	-.2844	-4.90E-02
	Equal variances not assumed			-2.817	76.969	.006	-.1667	5.917E-02	-.2845	-4.88E-02
Interface Quality	Equal variances assumed	1.573	.213	-7.755	82	.000	-.3631	4.682E-02	-.4562	-.2699
	Equal variances not assumed			-7.755	81.992	.000	-.3631	4.682E-02	-.4562	-.2699
Learn About Data	Equal variances assumed	4.838	.031	-4.378	82	.000	-.2321	5.302E-02	-.3376	-.1267
	Equal variances not assumed			-4.378	79.482	.000	-.2321	5.302E-02	-.3377	-.1266
Enjoyable	Equal variances assumed	.078	.781	-6.869	82	.000	-.3988	5.806E-02	-.5143	-.2833
	Equal variances not assumed			-6.869	81.989	.000	-.3988	5.806E-02	-.5143	-.2833
Usable	Equal variances assumed	6.480	.013	-3.072	82	.003	-.1726	5.619E-02	-.2844	-6.08E-02
	Equal variances not assumed			-3.072	77.608	.003	-.1726	5.619E-02	-.2845	-6.08E-02
Usability Survey Sum	Equal variances assumed	2.807	.098	-6.566	82	.000	-2.3274	.3545	-3.0325	-1.6223
	Equal variances not assumed			-6.566	80.778	.000	-2.3274	.3545	-3.0327	-1.6221

Table 46 T-Test – Satisfaction Survey

L.4 By Type

L.4.1 Descriptive Statistics

Group Statistics

	Treatment	N	Mean	Std. Deviation	Std. Error Mean
AnswerT1	JIM	42	2.71	.60	9.20E-02
	VIT	42	2.81	.51	7.80E-02
ConfT1Mean	JIM	42	.8958	.1552	2.394E-02
	VIT	42	.9177	.1601	2.470E-02
UserT1Mean	JIM	42	61.6270	30.3142	4.6776
	VIT	42	27.7421	14.0087	2.1616
AnswerT2	JIM	42	2.71	.55	8.54E-02
	VIT	42	2.88	.40	6.10E-02
ConfT2Mean	JIM	42	.8740	.2283	3.523E-02
	VIT	42	.9345	.1900	2.931E-02
UserT2Mean	JIM	42	95.3611	32.1651	4.9632
	VIT	42	29.3214	9.8772	1.5241
AnswersT3	JIM	42	2.43	.80	.12
	VIT	42	2.90	.30	4.58E-02
ConfT3Mean	JIM	42	.8433	.2553	3.940E-02
	VIT	42	.9444	.1068	1.648E-02
UserT3Mean	JIM	42	92.6746	37.9595	5.8573
	VIT	42	45.9048	12.8866	1.9885

Table 47 Descriptive Statistics – By Type

L.4.2 Hypothesis Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AnswerT1	Equal variances assumed	2.133	.148	-.790	82	.432	-9.52E-02	.12	-.34	.14
	Equal variances not assumed			-.790	79.865	.432	-9.52E-02	.12	-.34	.14
ConfT1Mean	Equal variances assumed	.360	.550	-.635	82	.528	-2.183E-02	3.440E-02	-9.03E-02	4.660E-02
	Equal variances not assumed			-.635	81.921	.528	-2.183E-02	3.440E-02	-9.03E-02	4.660E-02
UserT1Mean	Equal variances assumed	31.183	.000	6.576	82	.000	33.8849	5.1529	23.6342	44.1356
	Equal variances not assumed			6.576	57.748	.000	33.8849	5.1529	23.5693	44.2005
AnswerT2	Equal variances assumed	9.189	.003	-1.588	82	.116	-.17	.10	-.38	4.22E-02
	Equal variances not assumed			-1.588	74.168	.117	-.17	.10	-.38	4.25E-02
ConfT2Mean	Equal variances assumed	1.540	.218	-1.320	82	.190	-6.052E-02	4.583E-02	-.1517	3.065E-02
	Equal variances not assumed			-1.320	79.377	.190	-6.052E-02	4.583E-02	-.1517	3.070E-02
UserT2Mean	Equal variances assumed	32.292	.000	12.720	82	.000	66.0397	5.1919	55.7113	76.3681
	Equal variances not assumed			12.720	48.664	.000	66.0397	5.1919	55.6043	76.4751
AnswersT3	Equal variances assumed	64.591	.000	-3.613	82	.001	-.48	.13	-.74	-.21
	Equal variances not assumed			-3.613	52.080	.001	-.48	.13	-.74	-.21
ConfT3Mean	Equal variances assumed	21.601	.000	-2.369	82	.020	-.1012	4.271E-02	-.1861	-1.62E-02
	Equal variances not assumed			-2.369	54.919	.021	-.1012	4.271E-02	-.1868	-1.56E-02
UserT3Mean	Equal variances assumed	24.019	.000	7.561	82	.000	46.7698	6.1856	34.4647	59.0750
	Equal variances not assumed			7.561	50.327	.000	46.7698	6.1856	34.3477	59.1920

Table 48 T-Test – By Type

L.5 By Type Washout

L.5.1 Descriptive Statistics

Group Statistics

	Treatment	N	Mean	Std. Deviation	Std. Error Mean
AnswerT1	JIM	23	2.61	.72	.15
	VIT	19	2.84	.37	8.59E-02
ConfT1Mean	JIM	23	.8750	.1574	3.282E-02
	VIT	19	.9298	.1396	3.203E-02
UserT1Mean	JIM	23	67.4203	30.7057	6.4026
	VIT	19	34.6404	16.9575	3.8903
AnswerT2	JIM	23	2.70	.56	.12
	VIT	19	2.95	.23	5.26E-02
ConfT2Mean	JIM	23	.8804	.1856	3.870E-02
	VIT	19	.9474	.1422	3.263E-02
UserT2Mean	JIM	23	99.3333	35.7380	7.4519
	VIT	19	32.3596	10.3825	2.3819
AnswersT3	JIM	23	2.43	.79	.16
	VIT	19	2.89	.32	7.23E-02
ConfT3Mean	JIM	23	.8007	.2843	5.927E-02
	VIT	19	.9561	9.769E-02	2.241E-02
UserT3Mean	JIM	23	95.1812	46.4605	9.6877
	VIT	19	48.5526	11.7310	2.6913

Table 49 Descriptive Statistics – By Type – Washout

L.5.2 Box-plots

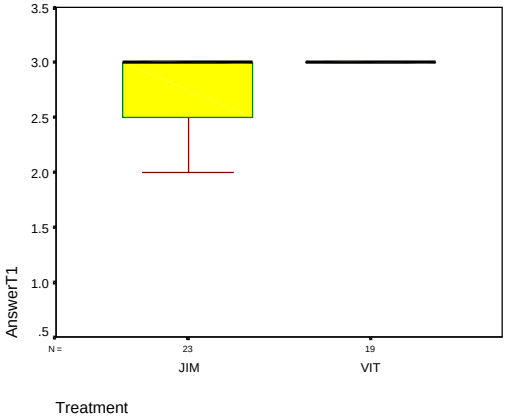


Figure 62 Correctness – Type I Washout Box-plot

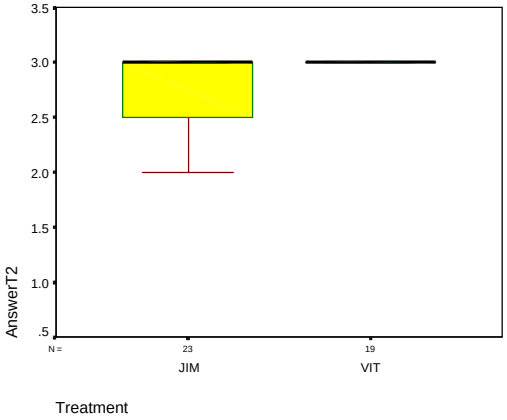


Figure 65 Correctness – Type II Washout Box-plot

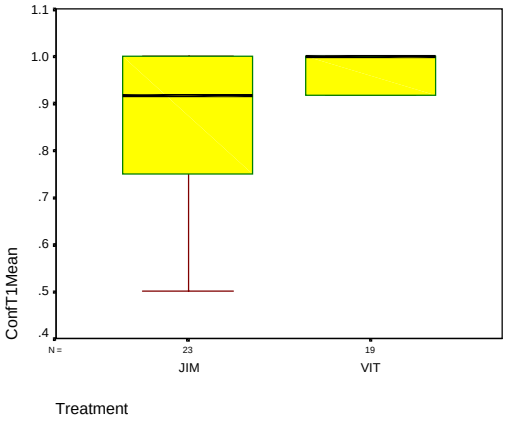


Figure 63 Confidence – Type I Washout Box-plot

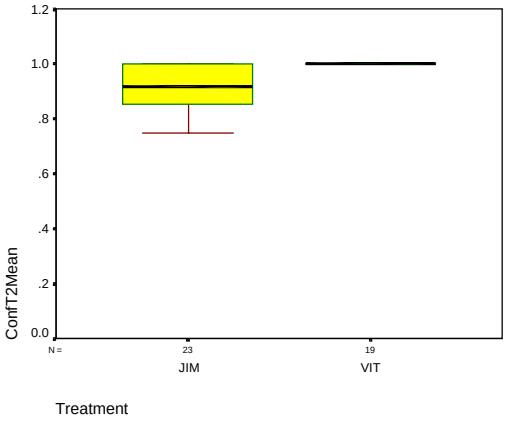


Figure 66 Confidence – Type II Washout Box-plot

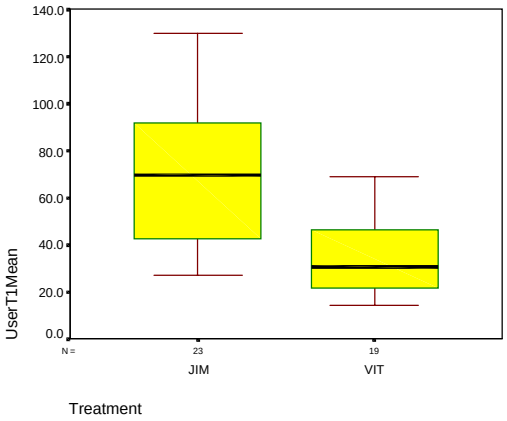


Figure 64 Speed – Type I Washout Box-plot

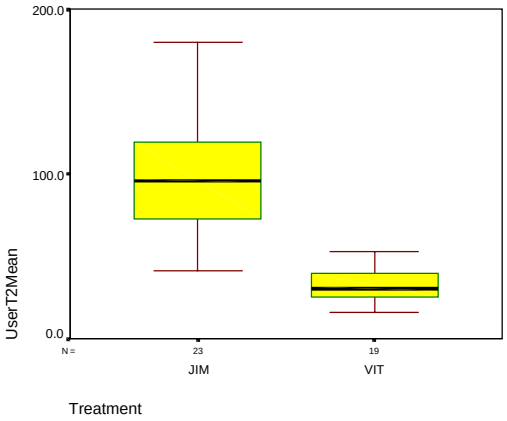


Figure 67 Speed – Type II Washout Box-plot

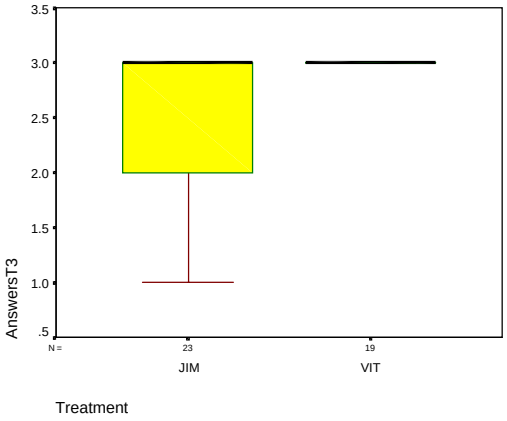


Figure 68 Correctness – Type III Washout Box-plot

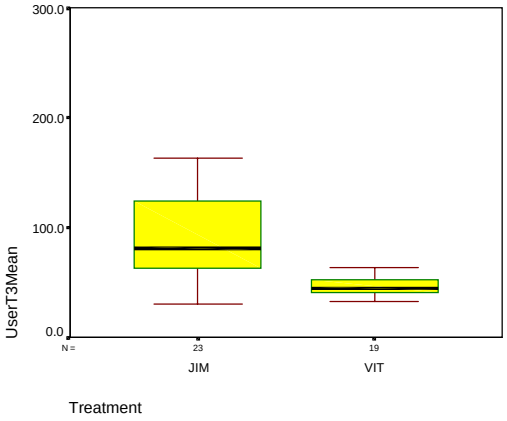


Figure 70 Speed – Type III Washout Box-plot

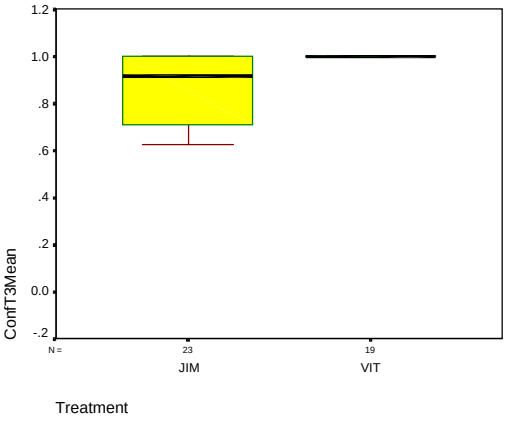


Figure 69 Confidence – Type III Washout Box-plot

L.5.3 Hypothesis Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AnswerT1	Equal variances assumed	8.187	.007	-1.272	40	.211	-.23	.18	-.60	.14
	Equal variances not assumed			-1.346	34.225	.187	-.23	.17	-.59	.12
ConfT1Mean	Equal variances assumed	1.958	.169	-1.182	40	.244	-5.482E-02	4.640E-02	-.1486	3.895E-02
	Equal variances not assumed			-1.195	39.771	.239	-5.482E-02	4.586E-02	-.1475	3.788E-02
UserT1Mean	Equal variances assumed	10.481	.002	4.154	40	.000	32.7799	7.8915	16.8306	48.7292
	Equal variances not assumed			4.375	35.354	.000	32.7799	7.4918	17.5761	47.9837
AnswerT2	Equal variances assumed	17.170	.000	-1.837	40	.074	-.25	.14	-.53	2.53E-02
	Equal variances not assumed			-1.969	30.349	.058	-.25	.13	-.51	9.27E-03
ConfT2Mean	Equal variances assumed	1.292	.263	-1.289	40	.205	-6.693E-02	5.192E-02	-.1719	3.801E-02
	Equal variances not assumed			-1.322	39.807	.194	-6.693E-02	5.062E-02	-.1693	3.539E-02
UserT2Mean	Equal variances assumed	15.397	.000	7.883	40	.000	66.9737	8.4956	49.8034	84.1440
	Equal variances not assumed			8.561	26.388	.000	66.9737	7.8233	50.9042	83.0432
AnswersT3	Equal variances assumed	26.803	.000	-2.388	40	.022	-.46	.19	-.85	-7.07E-02
	Equal variances not assumed			-2.563	29.982	.016	-.46	.18	-.83	-9.34E-02
ConfT3Mean	Equal variances assumed	10.506	.002	-2.271	40	.029	-.1554	6.844E-02	-.2937	-1.71E-02
	Equal variances not assumed			-2.453	28.040	.021	-.1554	6.337E-02	-.2852	-2.56E-02
UserT3Mean	Equal variances assumed	17.901	.000	4.256	40	.000	46.6285	10.9570	24.4837	68.7734
	Equal variances not assumed			4.638	25.342	.000	46.6285	10.0546	25.9349	67.3221

Table 50 T-Test – By Type - Washout

L.6 By Question

L.6.1 Descriptive Statistics

Group Statistics

	Treatment	N	Mean	Std. Deviation	Std. Error Mean
AnswersQ4	JIM	42	.86	.35	5.46E-02
	VIT	42	.88	.33	5.06E-02
ConfidenceQ4	JIM	42	.8333	.3007	4.641E-02
	VIT	42	.9107	.2126	3.280E-02
UserTimeQ4	JIM	42	91.36	54.33	8.38
	VIT	42	39.00	29.09	4.49
AnsweresQ5	JIM	42	1.00	.00[a]	.00
	VIT	42	1.00	.00[a]	.00
ConfidenceQ5	JIM	42	.8929	.1759	2.713E-02
	VIT	42	.8988	.1751	2.702E-02
UserTimeQ5	JIM	42	52.05	52.15	8.05
	VIT	42	24.33	13.68	2.11
AnswersQ6	JIM	42	.86	.35	5.46E-02
	VIT	42	.93	.26	4.02E-02
ConfidenceQ6	JIM	42	.8214	.2937	4.531E-02
	VIT	42	.9226	.2173	3.353E-02
UserTimeQ6	JIM	42	52.31	44.86	6.92
	VIT	42	20.14	10.12	1.56
AnswersQ7	JIM	42	.90	.30	4.58E-02
	VIT	42	1.00	.00	.00
ConfidenceQ7	JIM	42	.8155	.3027	4.671E-02
	VIT	42	.9405	.1897	2.928E-02
UserTimeQ7	JIM	42	137.67	51.34	7.92
	VIT	42	34.24	13.52	2.09
AnswersQ8	JIM	42	.93	.26	4.02E-02
	VIT	42	.98	.15	2.38E-02
ConfidenceQ8	JIM	42	.8452	.2867	4.423E-02
	VIT	42	.9345	.1836	2.833E-02
UserTimeQ8	JIM	42	78.14	53.58	8.27
	VIT	42	25.88	11.39	1.76
AnswersQ9	JIM	42	.88	.33	5.06E-02
	VIT	42	.90	.30	4.58E-02
ConfidenceQ9	JIM	42	.8690	.2225	3.433E-02
	VIT	42	.9464	.1299	2.005E-02
UserTimeQ9	JIM	42	76.21	31.26	4.82
	VIT	42	29.74	16.73	2.58
AnswersQ10	JIM	42	.95	.22	3.33E-02
	VIT	42	.93	.26	4.02E-02
ConfidenceQ10	JIM	42	.8929	.2416	3.728E-02
	VIT	42	.9405	.1078	1.663E-02
UserTimeQ10	JIM	42	69.81	39.06	6.03
	VIT	42	70.21	21.87	3.37
AnswersQ11	JIM	42	.76	.43	6.65E-02
	VIT	42	1.00	.00	.00
ConfidenceQ11	JIM	42	.7024	.3589	5.537E-02
	VIT	42	.9345	.1568	2.419E-02
UserTimeQ11	JIM	42	134.90	77.81	12.01
	VIT	42	33.64	11.91	1.84
AnswersQ12	JIM	42	.71	.46	7.06E-02
	VIT	42	.98	.15	2.38E-02
ConfidenceQ12	JIM	42	.7321	.3376	5.210E-02
	VIT	42	.9524	.1379	2.128E-02
UserTimeQ12	JIM	42	73.10	43.46	6.71
	VIT	42	35.74	10.86	1.68

a. t cannot be computed because the standard deviations of both groups are 0.

Table 51 Descriptive Statistics – By Question

L.6.2 Box-plots

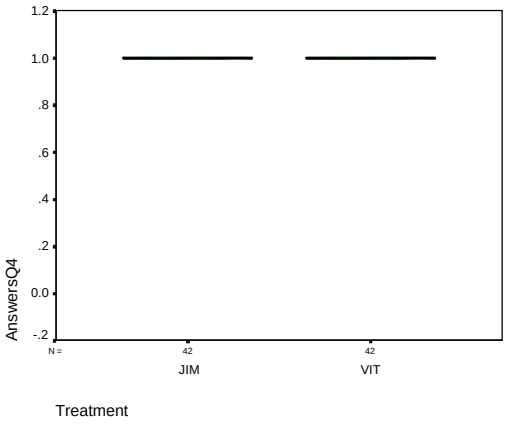


Figure 71 Correctness – Task 4 Box-plot

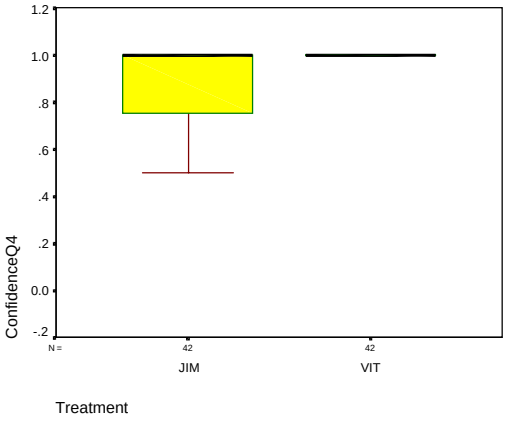


Figure 72 Confidence – Task 4 Box-plot

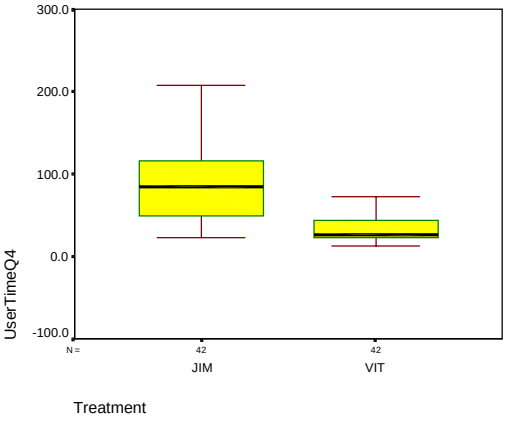


Figure 73 Speed – Task 4 Box-plot

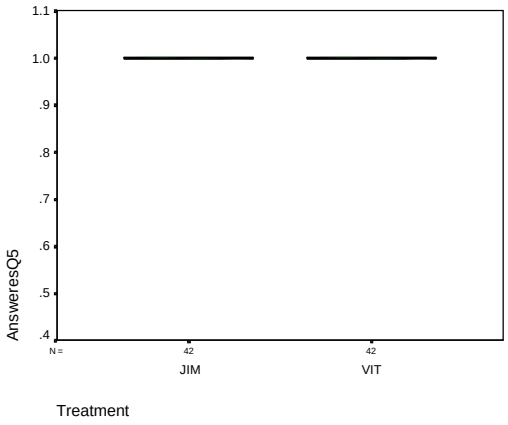


Figure 74 Correctness – Task 5 Box-plot

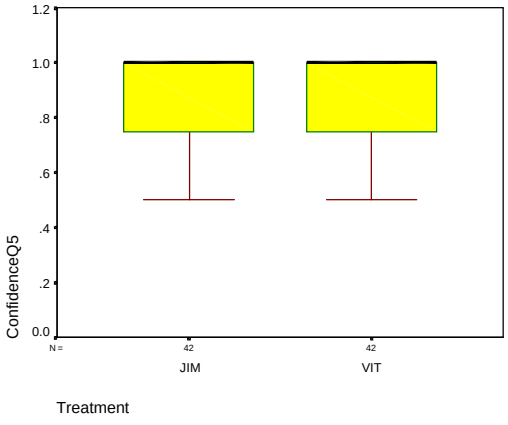


Figure 75 Confidence – Task 5 Box-plot

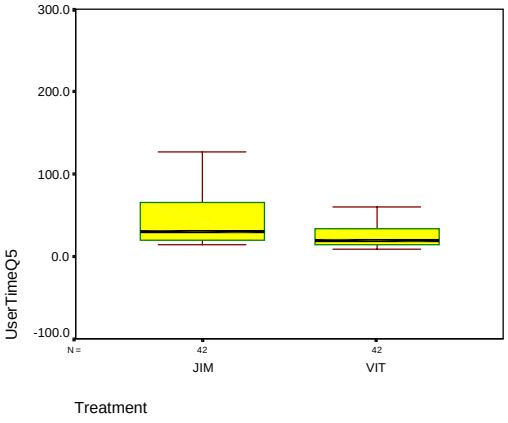


Figure 76 Speed – Task 5 Box-plot

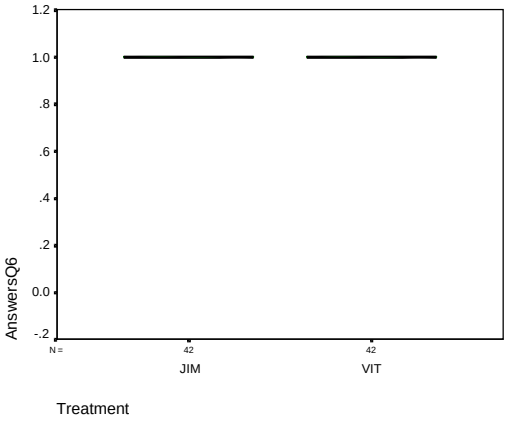


Figure 77 Correctness – Task 6 Box-plot

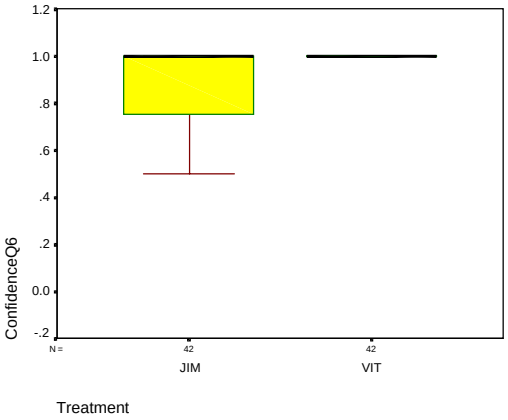


Figure 78 Confidence – Task 6 Box-plot

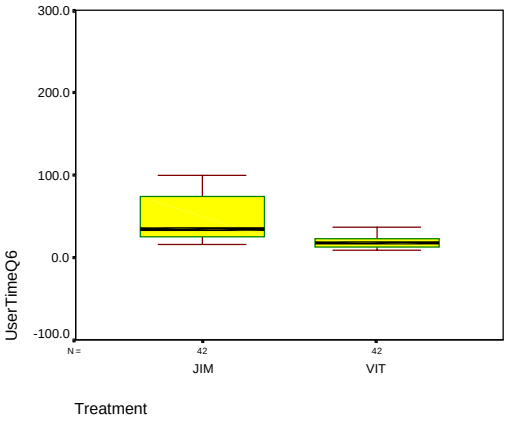


Figure 79 Speed – Task 6 Box-plot

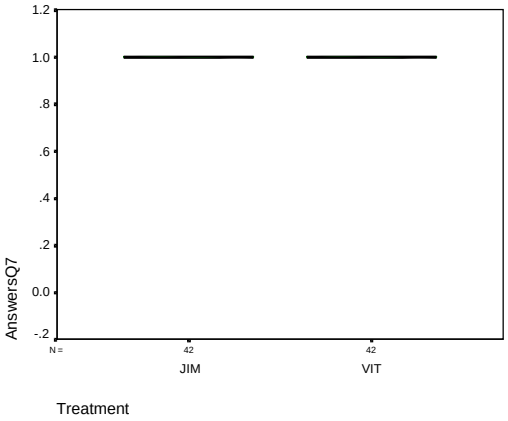


Figure 80 Correctness – Task 7 Box-plot

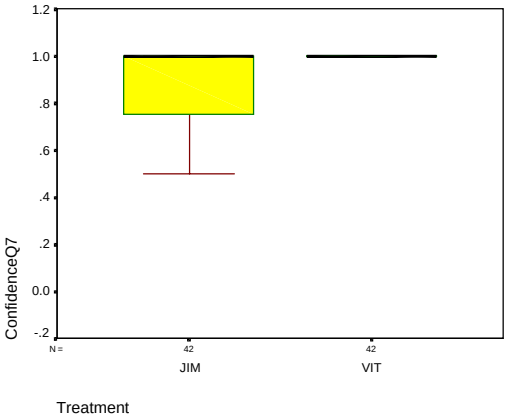


Figure 81 Confidence – Task 7 Box-plot

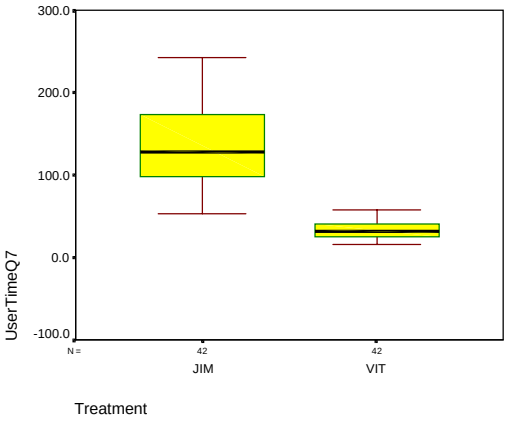


Figure 82 Speed – Task 7 Box-plot

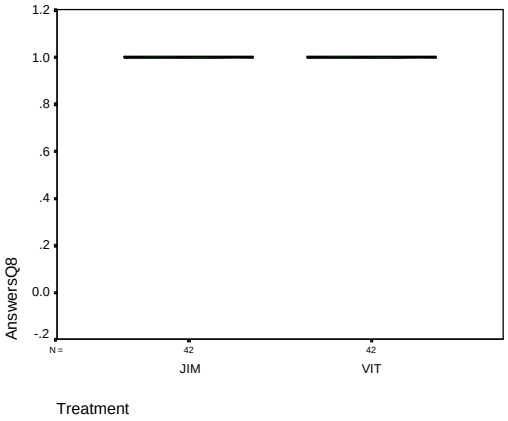


Figure 83 Correctness – Task 8 Box-plot

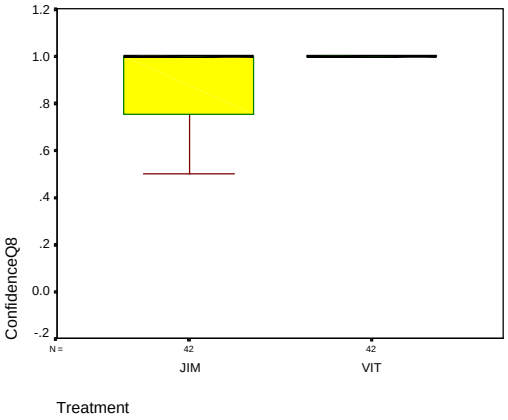


Figure 84 Confidence – Task 8 Box-plot

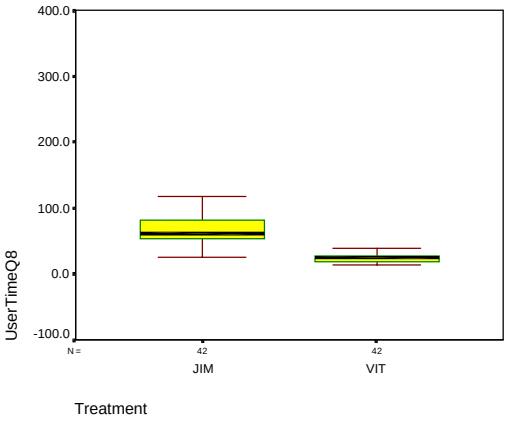


Figure 85 Speed – Task 8 Box-plot

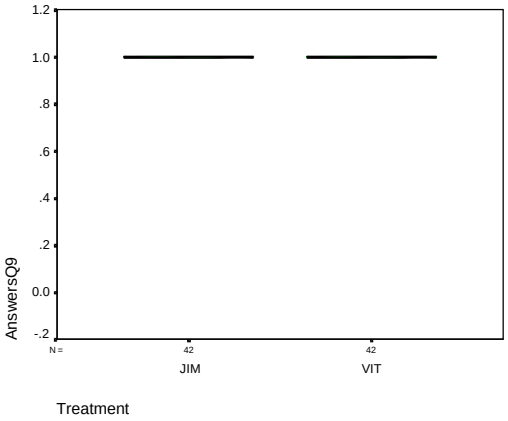


Figure 86 Correctness – Task 9 Box-plot

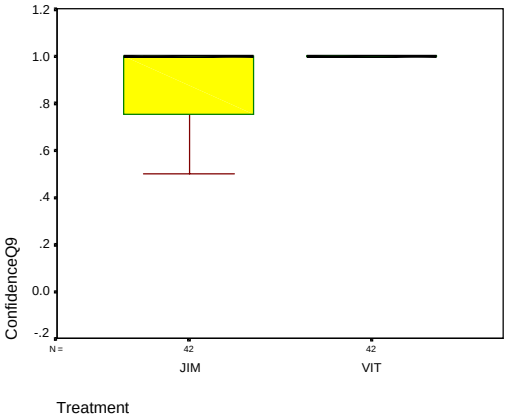


Figure 87 Confidence – Task 9 Box-plot

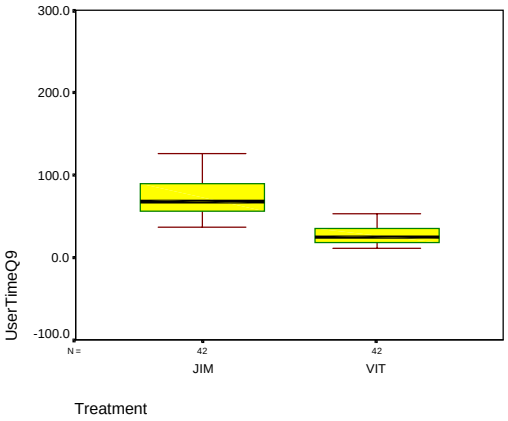


Figure 88 Speed – Task 9 Box-plot

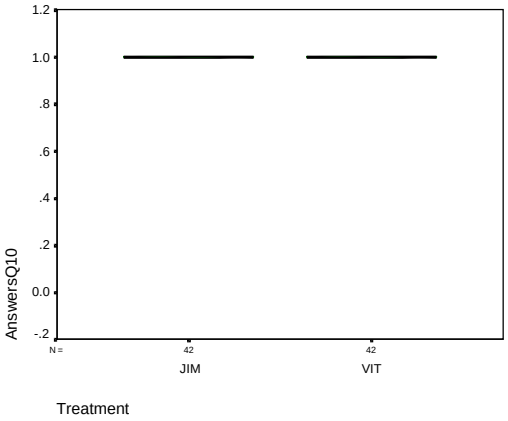


Figure 89 Correctness – Task 10 Box-plot

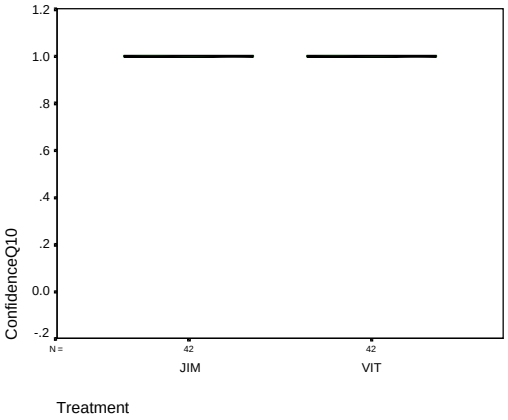


Figure 90 Confidence – Task 10 Box-plot

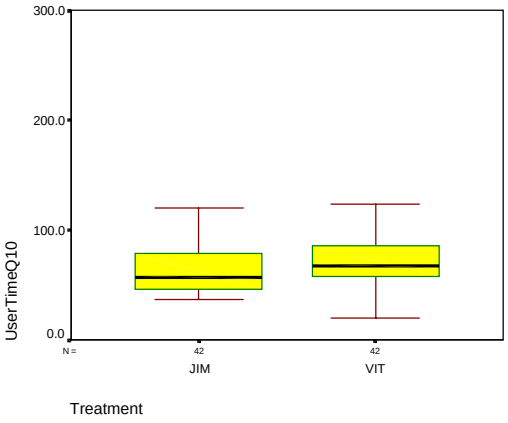


Figure 91 Speed – Task 10 Box-plot

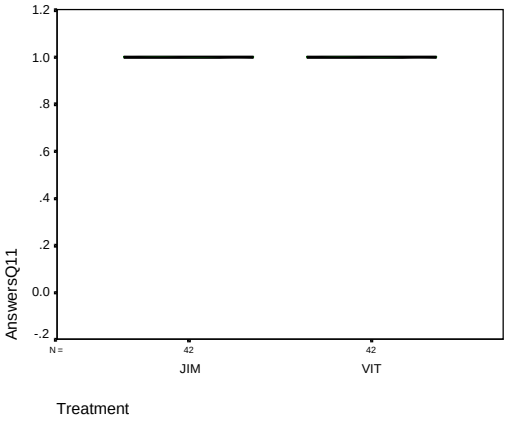


Figure 92 Correctness – Task 11 Box-plot

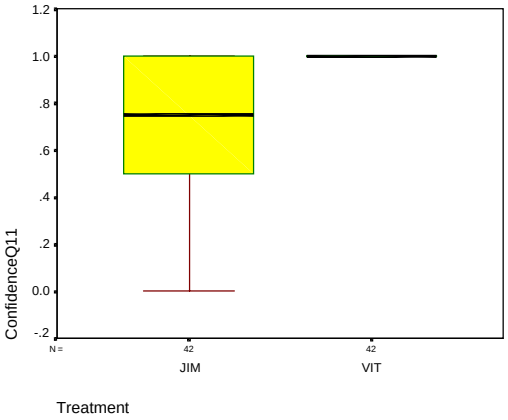


Figure 93 Confidence – Task 11 Box-plot

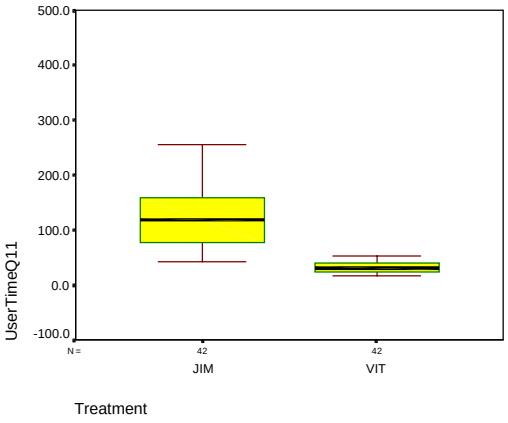


Figure 94 Speed – Task 11 Box-plot

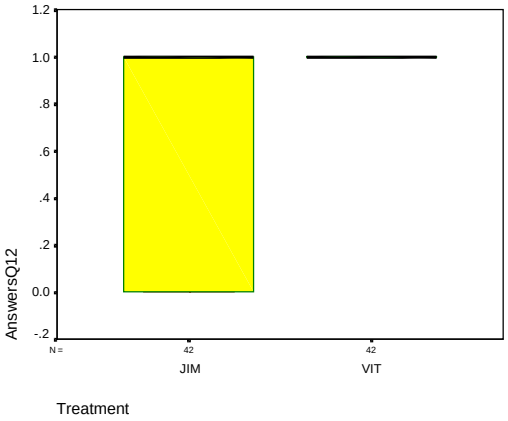


Figure 95 Correctness – Task 12 Box-plot

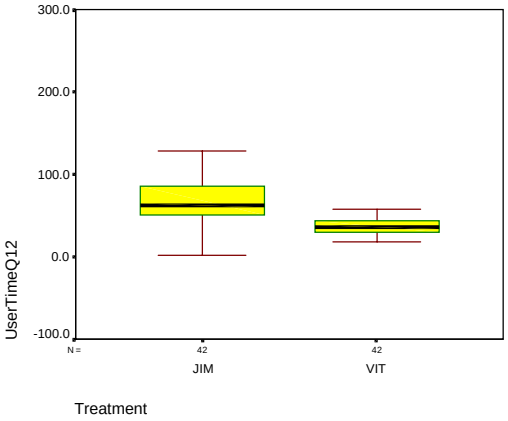


Figure 97 Speed – Task 12 Box-plot

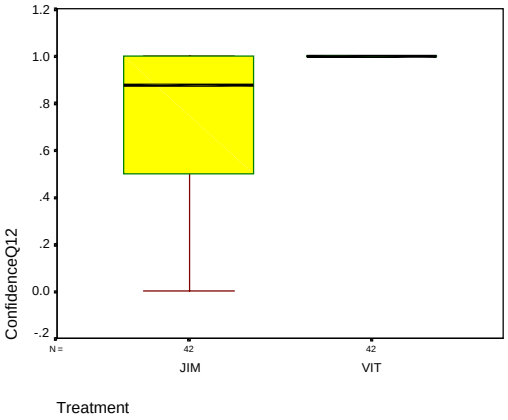


Figure 96 Confidence – Task 12 Box-plot

L.6.3 Hypothesis Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AnswersQ4	Equal variances assumed	.411	.523	-.320	82	.750	-2.38E-02	7.45E-02	-.17	.12
	Equal variances not assumed			-.320	81.513	.750	-2.38E-02	7.45E-02	-.17	.12
ConfidenceQ4	Equal variances assumed	3.981	.049	-1.362	82	.177	-7.738E-02	5.683E-02	-.1904	3.567E-02
	Equal variances not assumed			-1.362	73.784	.177	-7.738E-02	5.683E-02	-.1906	3.586E-02
UserTimeQ4	Equal variances assumed	9.804	.002	5.506	82	.000	52.36	9.51	33.44	71.27
	Equal variances not assumed			5.506	62.727	.000	52.36	9.51	33.35	71.36
ConfidenceQ5	Equal variances assumed	.006	.937	-.155	82	.877	-5.952E-03	3.829E-02	-8.21E-02	7.023E-02
	Equal variances not assumed			-.155	81.999	.877	-5.952E-03	3.829E-02	-8.21E-02	7.023E-02
UserTimeQ5	Equal variances assumed	20.477	.000	3.332	82	.001	27.71	8.32	11.17	44.26
	Equal variances not assumed			3.332	46.614	.002	27.71	8.32	10.98	44.45
AnswersQ6	Equal variances assumed	4.645	.034	-1.053	82	.296	-7.14E-02	6.79E-02	-.21	6.36E-02
	Equal variances not assumed			-1.053	75.341	.296	-7.14E-02	6.79E-02	-.21	6.37E-02
ConfidenceQ6	Equal variances assumed	5.026	.028	-1.795	82	.076	-.1012	5.637E-02	-.2133	1.095E-02
	Equal variances not assumed			-1.795	75.543	.077	-.1012	5.637E-02	-.2135	1.109E-02
UserTimeQ6	Equal variances assumed	20.219	.000	4.533	82	.000	32.17	7.10	18.05	46.28
	Equal variances not assumed			4.533	45.164	.000	32.17	7.10	17.87	46.46
AnswersQ7	Equal variances assumed	21.564	.000	-2.077	82	.041	-9.52E-02	4.58E-02	-.19	-4.04E-03
	Equal variances not assumed			-2.077	41.000	.044	-9.52E-02	4.58E-02	-.19	-2.65E-03
ConfidenceQ7	Equal variances assumed	10.195	.002	-2.267	82	.026	-.1250	5.513E-02	-.2347	-1.53E-02
	Equal variances not assumed			-2.267	68.908	.027	-.1250	5.513E-02	-.2350	-1.50E-02
UserTimeQ7	Equal variances assumed	46.114	.000	12.624	82	.000	103.43	8.19	87.13	119.73
	Equal variances not assumed			12.624	46.663	.000	103.43	8.19	86.94	119.91
AnswersQ8	Equal variances assumed	4.361	.040	-1.019	82	.311	-4.76E-02	4.67E-02	-.14	4.54E-02
	Equal variances not assumed			-1.019	66.592	.312	-4.76E-02	4.67E-02	-.14	4.57E-02
ConfidenceQ8	Equal variances assumed	7.696	.007	-1.700	82	.093	-8.929E-02	5.253E-02	-.1938	1.521E-02
	Equal variances not assumed			-1.700	69.799	.094	-8.929E-02	5.253E-02	-.1941	1.549E-02
UserTimeQ8	Equal variances assumed	15.135	.000	6.184	82	.000	52.26	8.45	35.45	69.07
	Equal variances not assumed			6.184	44.695	.000	52.26	8.45	35.24	69.29
AnswersQ9	Equal variances assumed	.489	.486	-.349	82	.728	-2.38E-02	6.83E-02	-.16	.11
	Equal variances not assumed			-.349	81.221	.728	-2.38E-02	6.83E-02	-.16	.11
ConfidenceQ9	Equal variances assumed	9.002	.004	-1.946	82	.055	-7.738E-02	3.976E-02	-.1565	1.707E-03
	Equal variances not assumed			-1.946	66.044	.056	-7.738E-02	3.976E-02	-.1568	1.994E-03
UserTimeQ9	Equal variances assumed	7.699	.007	8.494	82	.000	46.48	5.47	35.59	57.36
	Equal variances not assumed			8.494	62.699	.000	46.48	5.47	35.54	57.41
AnswersQ10	Equal variances assumed	.840	.362	.456	82	.649	2.38E-02	5.22E-02	-8.00E-02	.13
	Equal variances not assumed			.456	79.206	.649	2.38E-02	5.22E-02	-8.01E-02	.13
ConfidenceQ10	Equal variances assumed	7.792	.007	-1.167	82	.247	-4.762E-02	4.082E-02	-.1288	3.358E-02
	Equal variances not assumed			-1.167	56.697	.248	-4.762E-02	4.082E-02	-.1294	3.413E-02
UserTimeQ10	Equal variances assumed	4.016	.048	-.059	82	.953	-.40	6.91	-14.15	13.34
	Equal variances not assumed			-.059	64.396	.953	-.40	6.91	-14.20	13.39
AnswersQ11	Equal variances assumed	108.430	.000	-3.579	82	.001	-.24	6.65E-02	-.37	-.11
	Equal variances not assumed			-3.579	41.000	.001	-.24	6.65E-02	-.37	-.10
ConfidenceQ11	Equal variances assumed	33.015	.000	-3.842	82	.000	-.2321	6.042E-02	-.3523	-.1119
	Equal variances not assumed			-3.842	56.097	.000	-.2321	6.042E-02	-.3532	-.1111
UserTimeQ11	Equal variances assumed	31.111	.000	8.337	82	.000	101.26	12.15	77.10	125.43
	Equal variances not assumed			8.337	42.921	.000	101.26	12.15	76.76	125.76
AnswersQ12	Equal variances assumed	91.576	.000	-3.517	82	.001	-.26	7.45E-02	-.41	-.11
	Equal variances not assumed			-3.517	50.219	.001	-.26	7.45E-02	-.41	-.11
ConfidenceQ12	Equal variances assumed	31.991	.000	-3.913	82	.000	-.2202	5.628E-02	-.3322	-.1083
	Equal variances not assumed			-3.913	54.309	.000	-.2202	5.628E-02	-.3331	-.1074
UserTimeQ12	Equal variances assumed	22.001	.000	5.405	82	.000	37.36	6.91	23.61	51.11
	Equal variances not assumed			5.405	46.102	.000	37.36	6.91	23.45	51.27

Table 52 T-Test – By Question

L.7 By Question Washout

L.7.1 Descriptive Statistics

Group Statistics

	Treatment	N	Mean	Std. Deviation	Std. Error Mean
AnswersQ4	JIM	23	.78	.42	8.79E-02
	VIT	19	.89	.32	7.23E-02
ConfidenceQ4	JIM	23	.7609	.3327	6.937E-02
	VIT	19	.9211	.1678	3.849E-02
UserTimeQ4	JIM	23	103.70	57.11	11.91
	VIT	19	53.16	37.38	8.57
AnsweresQ5	JIM	23	1.00	.00[a]	.00
	VIT	19	1.00	.00[a]	.00
ConfidenceQ5	JIM	23	.9022	.1458	3.039E-02
	VIT	19	.9079	.1710	3.923E-02
UserTimeQ5	JIM	23	52.70	48.92	10.20
	VIT	19	28.84	14.29	3.28
AnswersQ6	JIM	23	.83	.39	8.08E-02
	VIT	19	.95	.23	5.26E-02
ConfidenceQ6	JIM	23	.7717	.3190	6.652E-02
	VIT	19	.9605	.1254	2.876E-02
UserTimeQ6	JIM	23	53.48	24.22	5.05
	VIT	19	23.95	10.97	2.52
AnswersQ7	JIM	23	.87	.34	7.18E-02
	VIT	19	1.00	.00	.00
ConfidenceQ7	JIM	23	.7717	.3190	6.652E-02
	VIT	19	.9605	.1254	2.876E-02
UserTimeQ7	JIM	23	133.65	50.75	10.58
	VIT	19	36.74	10.76	2.47
AnswersQ8	JIM	23	.91	.29	6.01E-02
	VIT	19	1.00	.00	.00
ConfidenceQ8	JIM	23	.8152	.2940	6.131E-02
	VIT	19	.9342	.1834	4.207E-02
UserTimeQ8	JIM	23	90.13	65.72	13.70
	VIT	19	26.53	11.03	2.53
AnswersQ9	JIM	23	.91	.29	6.01E-02
	VIT	19	.95	.23	5.26E-02
ConfidenceQ9	JIM	23	.8804	.1663	3.468E-02
	VIT	19	.9211	.1678	3.849E-02
UserTimeQ9	JIM	23	81.65	37.86	7.89
	VIT	19	35.42	21.17	4.86
AnswersQ10	JIM	23	.96	.21	4.35E-02
	VIT	19	.89	.32	7.23E-02
ConfidenceQ10	JIM	23	.8804	.2485	5.182E-02
	VIT	19	.9737	7.883E-02	1.808E-02
UserTimeQ10	JIM	23	65.09	43.38	9.04
	VIT	19	75.42	15.79	3.62
AnswersQ11	JIM	23	.78	.42	8.79E-02
	VIT	19	1.00	.00	.00
ConfidenceQ11	JIM	23	.6630	.3739	7.797E-02
	VIT	19	.9342	.1405	3.223E-02
UserTimeQ11	JIM	23	153.57	95.06	19.82
	VIT	19	36.74	13.93	3.20
AnswersQ12	JIM	23	.70	.47	9.81E-02
	VIT	19	1.00	.00	.00
ConfidenceQ12	JIM	23	.7500	.3015	6.287E-02
	VIT	19	.9605	9.366E-02	2.149E-02
UserTimeQ12	JIM	23	77.70	38.91	8.11
	VIT	19	37.16	13.63	3.13

a. t cannot be computed because the standard deviations of both groups are 0.

Table 53 Descriptive Statistics – By Question – Washout

L.7.2 Box-plots

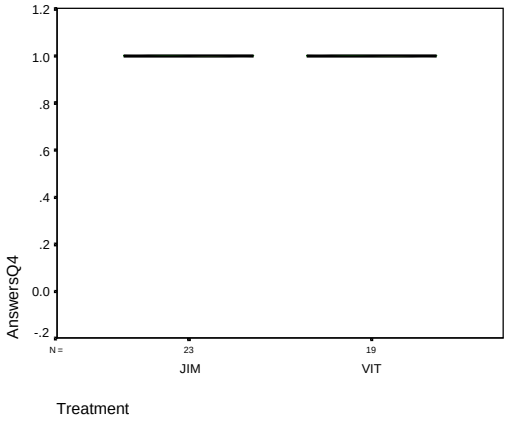


Figure 98 Correctness – Task 4 Washout Box-plot

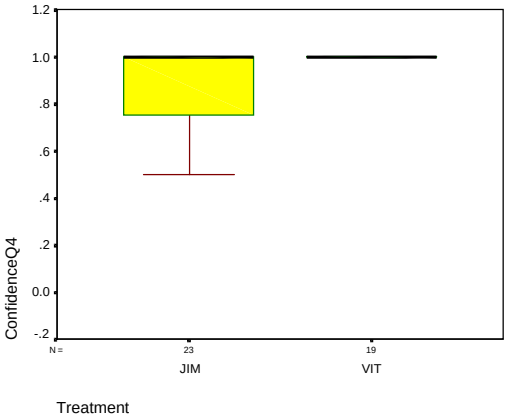


Figure 99 Confidence – Task 4 Washout Box-plot

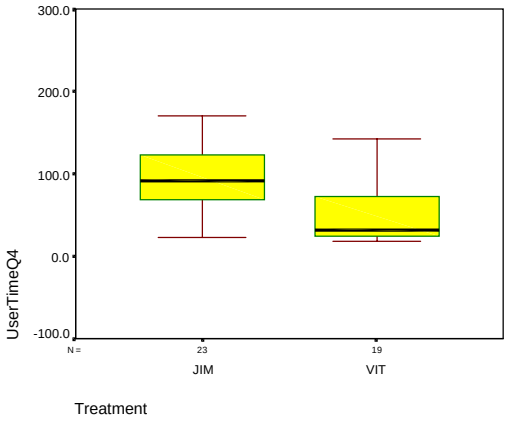


Figure 100 Speed – Task 4 Washout Box-plot

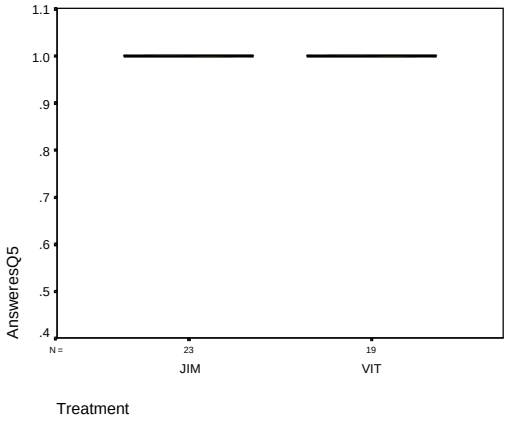


Figure 101 Correctness – Task 5 Washout Box-plot

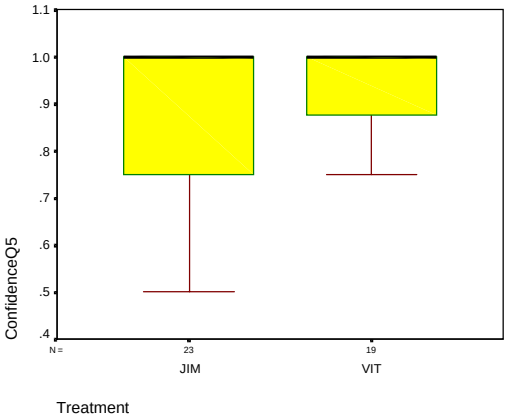


Figure 102 Confidence – Task 5 Washout Box-plot

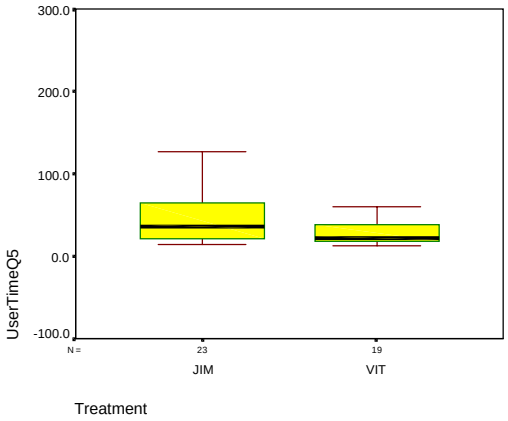


Figure 103 Speed – Task 5 Washout Box-plot

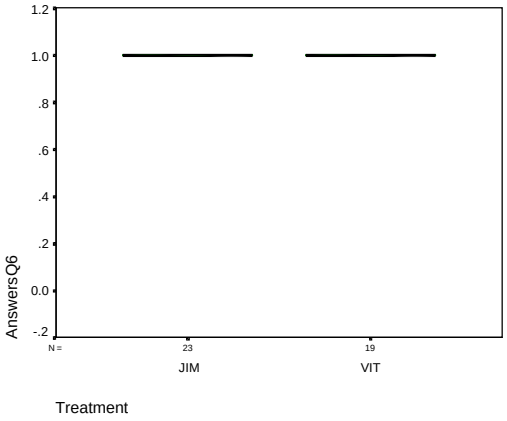


Figure 104 Correctness – Task 6 Washout Box-plot

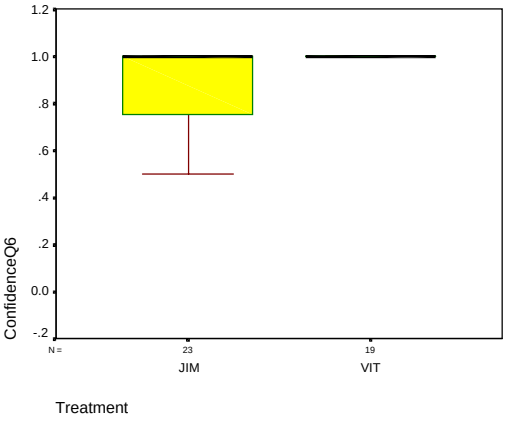


Figure 105 Confidence – Task 6 Washout Box-plot

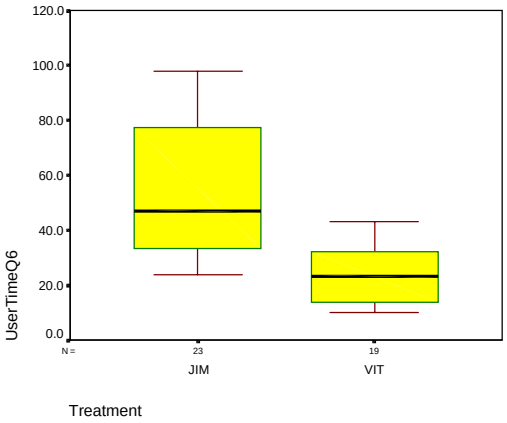


Figure 106 Speed – Task 6 Washout Box-plot

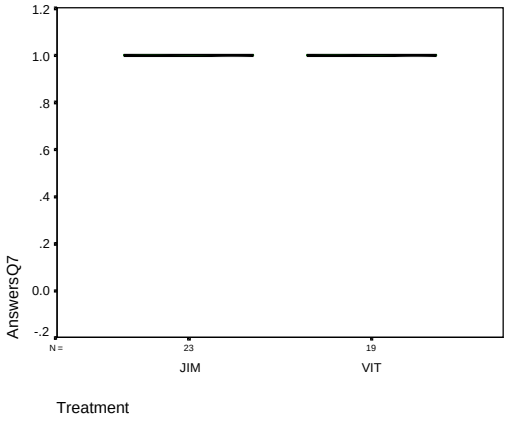


Figure 107 Correctness – Task 7 Washout Box-plot

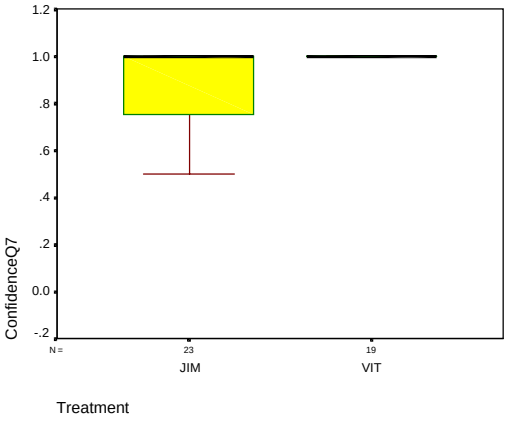


Figure 108 Confidence – Task 7 Washout Box-plot

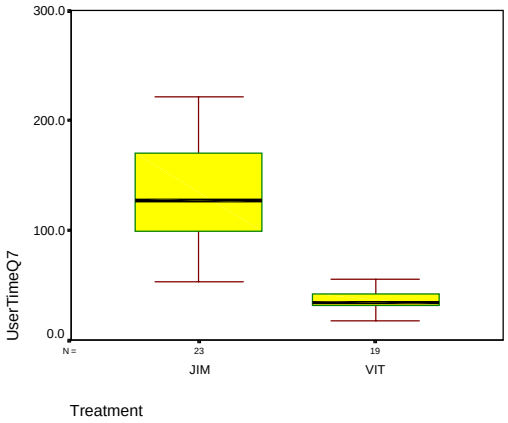


Figure 109 Speed – Task 7 Washout Box-plot

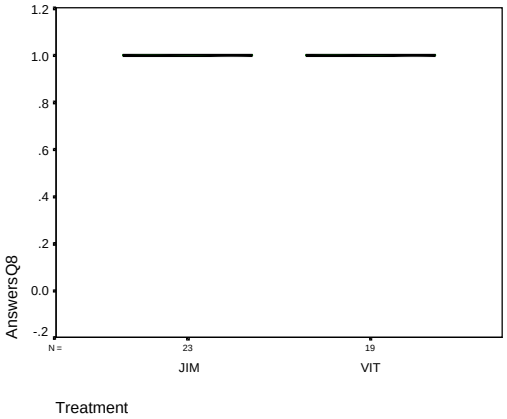


Figure 110 Correctness – Task 8 Washout Box-plot

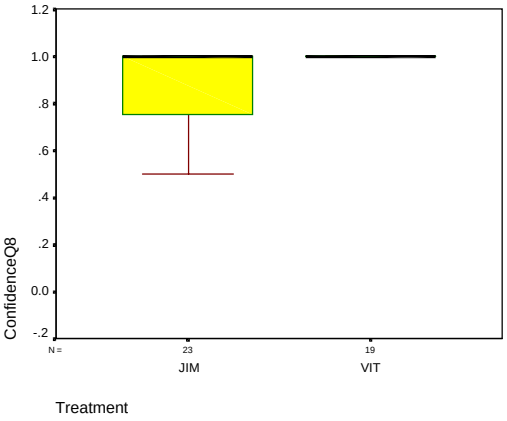


Figure 111 Confidence – Task 8 Washout Box-plot

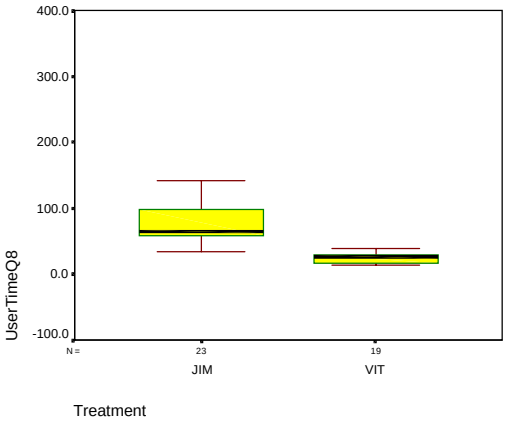


Figure 112 Speed – Task 8 Washout Box-plot

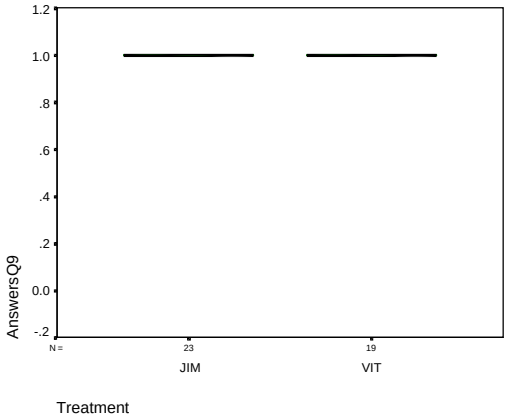


Figure 113 Correctness – Task 9 Washout Box-plot

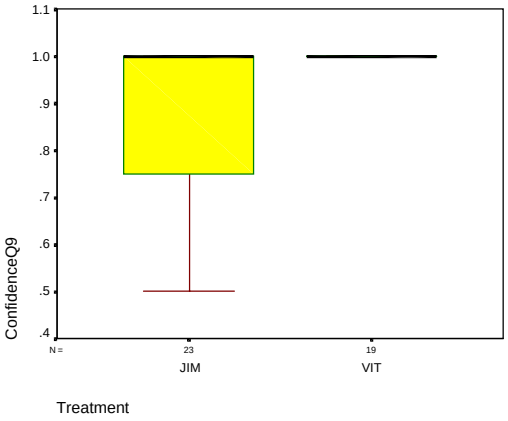


Figure 114 Confidence – Task 9 Washout Box-plot

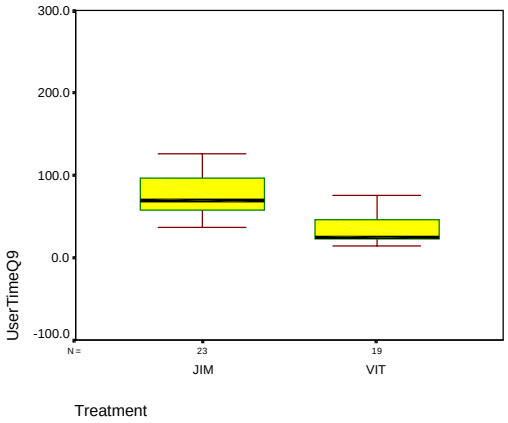


Figure 115 Speed – Task 9 Washout Box-plot

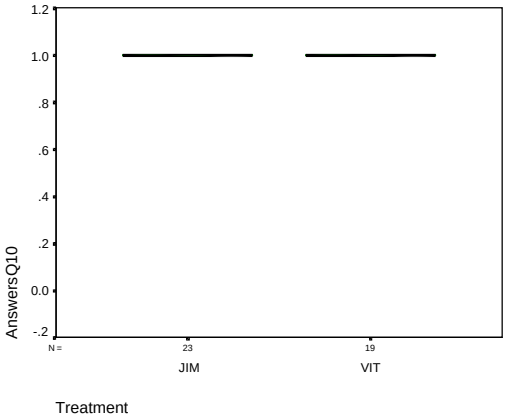


Figure 116 Correctness – Task 10 Washout Box-plot

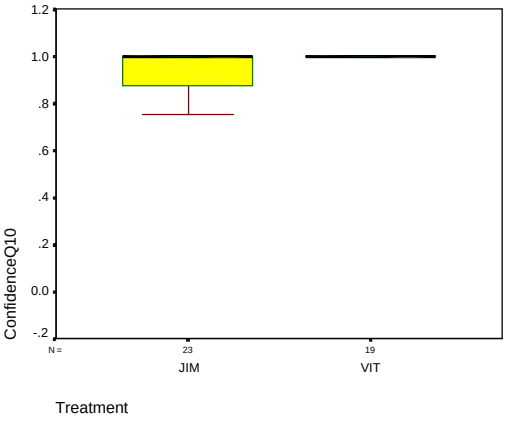


Figure 117 Confidence – Task 10 Washout Box-plot

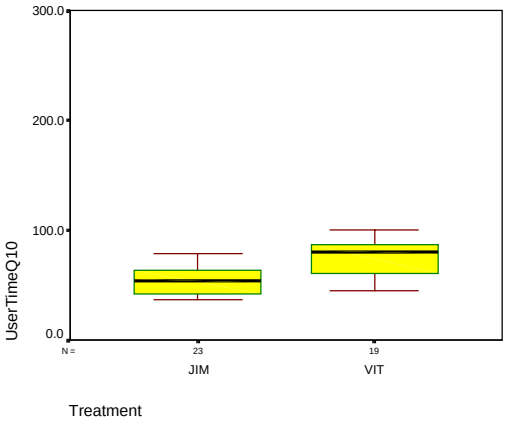


Figure 118 Speed – Task 10 Washout Box-plot

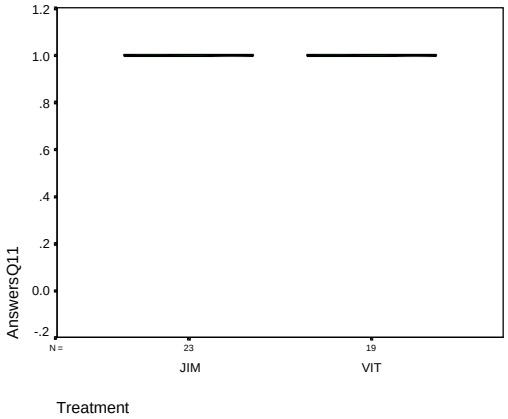


Figure 119 Correctness – Task 11 Washout Box-plot

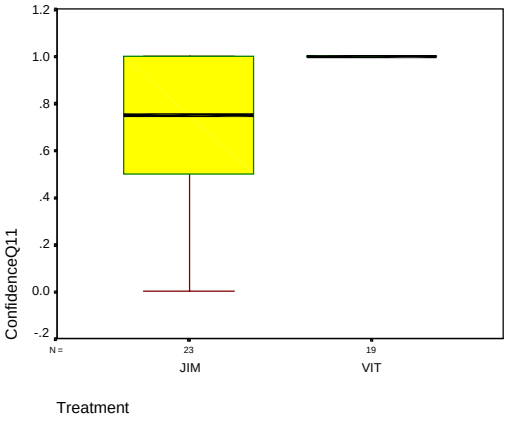


Figure 120 Confidence – Task 11 Washout Box-plot

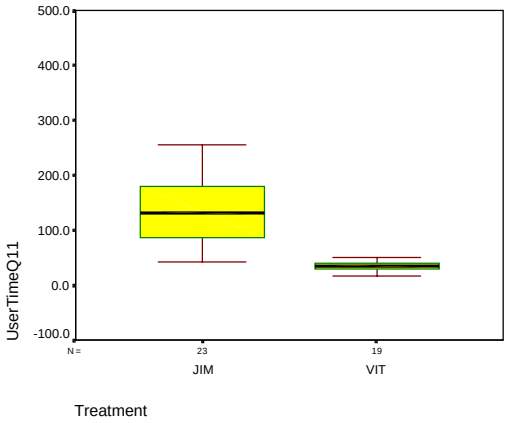


Figure 121 Speed – Task 11 Washout Box-plot

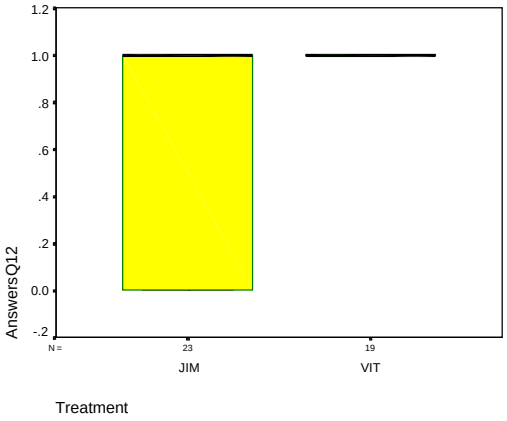


Figure 122 Correctness – Task 12 Washout Box-plot

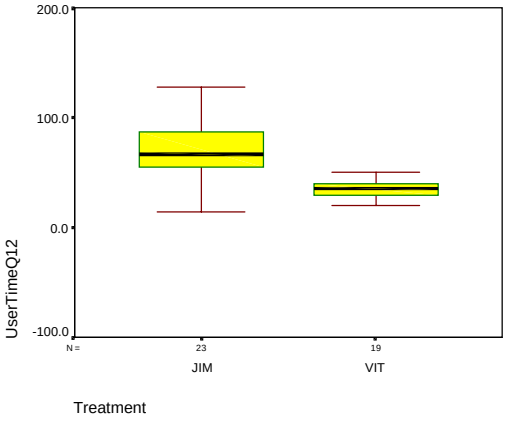


Figure 124 Speed – Task 12 Washout Box-plot

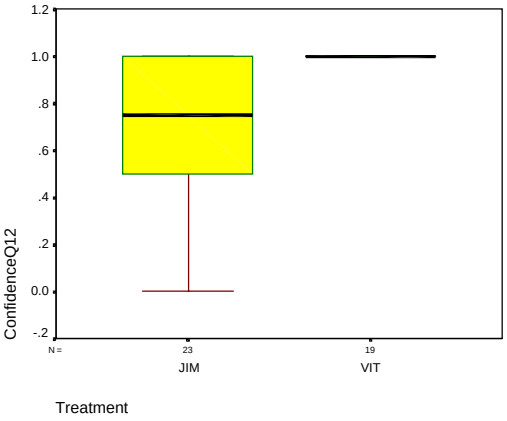


Figure 123 Confidence – Task 12 Washout Box-plot

L.7.3 Hypothesis Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AnswersQ4	Equal variances assumed	4.060	.051	-.958	40	.344	-.11	.12	-.35	.12
	Equal variances not assumed			-.985	39.654	.331	-.11	.11	-.34	.12
ConfidenceQ4	Equal variances assumed	5.346	.026	-1.905	40	.064	-.1602	8.407E-02	-.3301	9.724E-03
	Equal variances not assumed			-2.019	33.723	.051	-.1602	7.933E-02	-.3214	1.081E-03
UserTimeQ4	Equal variances assumed	1.070	.307	3.312	40	.002	50.54	15.26	19.70	81.38
	Equal variances not assumed			3.444	38.183	.001	50.54	14.67	20.84	80.24
ConfidenceQ5	Equal variances assumed	.102	.751	-.117	40	.907	-5.72E-03	4.886E-02	-.1045	9.304E-02
	Equal variances not assumed			-.115	35.599	.909	-5.72E-03	4.962E-02	-.1064	9.496E-02
UserTimeQ5	Equal variances assumed	6.834	.013	2.050	40	.047	23.85	11.63	.34	47.37
	Equal variances not assumed			2.226	26.433	.035	23.85	10.71	1.85	45.86
AnswersQ6	Equal variances assumed	6.770	.013	-1.200	40	.237	-.12	.10	-.33	8.30E-02
	Equal variances not assumed			-1.258	36.578	.217	-.12	9.64E-02	-.32	7.42E-02
ConfidenceQ6	Equal variances assumed	10.410	.003	-2.425	40	.020	-.1888	7.784E-02	-.3461	-3.15E-02
	Equal variances not assumed			-2.605	29.724	.014	-.1888	7.247E-02	-.3369	-4.07E-02
UserTimeQ6	Equal variances assumed	21.871	.000	4.907	40	.000	29.53	6.02	17.37	41.69
	Equal variances not assumed			5.233	31.882	.000	29.53	5.64	18.03	41.03
AnswersQ7	Equal variances assumed	15.027	.000	-1.648	40	.107	-.13	7.92E-02	-.29	2.96E-02
	Equal variances not assumed			-1.817	22.000	.083	-.13	7.18E-02	-.28	1.85E-02
ConfidenceQ7	Equal variances assumed	10.410	.003	-2.425	40	.020	-.1888	7.784E-02	-.3461	-3.15E-02
	Equal variances not assumed			-2.605	29.724	.014	-.1888	7.247E-02	-.3369	-4.07E-02
UserTimeQ7	Equal variances assumed	19.261	.000	8.157	40	.000	96.92	11.88	72.90	120.93
	Equal variances not assumed			8.919	24.371	.000	96.92	10.87	74.51	119.32
AnswersQ8	Equal variances assumed	8.421	.006	-1.313	40	.197	-8.70E-02	6.62E-02	-.22	4.69E-02
	Equal variances not assumed			-1.447	22.000	.162	-8.70E-02	6.01E-02	-.21	3.76E-02
ConfidenceQ8	Equal variances assumed	4.884	.033	-1.533	40	.133	-.1190	7.762E-02	-.2759	3.788E-02
	Equal variances not assumed			-1.600	37.447	.118	-.1190	7.436E-02	-.2696	3.161E-02
UserTimeQ8	Equal variances assumed	9.241	.004	4.162	40	.000	63.60	15.28	32.72	94.49
	Equal variances not assumed			4.564	23.492	.000	63.60	13.93	34.81	92.40
AnswersQ9	Equal variances assumed	.724	.400	-.420	40	.676	-3.43E-02	8.16E-02	-.20	.13
	Equal variances not assumed			-.430	39.960	.670	-3.43E-02	7.99E-02	-.20	.13
ConfidenceQ9	Equal variances assumed	.547	.464	-.785	40	.437	-4.06E-02	5.177E-02	-.1452	6.401E-02
	Equal variances not assumed			-.784	38.391	.438	-4.06E-02	5.181E-02	-.1455	6.423E-02
UserTimeQ9	Equal variances assumed	2.804	.102	4.740	40	.000	46.23	9.75	26.52	65.94
	Equal variances not assumed			4.988	35.575	.000	46.23	9.27	27.43	65.03
AnswersQ10	Equal variances assumed	2.408	.129	.761	40	.451	6.18E-02	8.12E-02	-.10	.23
	Equal variances not assumed			.732	30.137	.470	6.18E-02	8.44E-02	-.11	.23
ConfidenceQ10	Equal variances assumed	9.856	.003	-1.569	40	.125	-9.32E-02	5.944E-02	-.2134	2.689E-02
	Equal variances not assumed			-1.699	27.192	.101	-9.32E-02	5.488E-02	-.2058	1.932E-02
UserTimeQ10	Equal variances assumed	2.191	.147	-.984	40	.331	-10.33	10.50	-31.55	10.89
	Equal variances not assumed			-1.061	28.719	.298	-10.33	9.74	-30.27	9.60
AnswersQ11	Equal variances assumed	38.546	.000	-2.242	40	.031	-.22	9.70E-02	-.41	-2.14E-02
	Equal variances not assumed			-2.472	22.000	.022	-.22	8.79E-02	-.40	-3.50E-02
ConfidenceQ11	Equal variances assumed	17.026	.000	-2.986	40	.005	-.2712	9.080E-02	-.4547	-8.77E-02
	Equal variances not assumed			-3.214	29.122	.003	-.2712	8.437E-02	-.4437	-9.86E-02
UserTimeQ11	Equal variances assumed	17.706	.000	5.299	40	.000	116.83	22.05	72.27	161.39
	Equal variances not assumed			5.819	23.139	.000	116.83	20.08	75.31	158.35
AnswersQ12	Equal variances assumed	100.082	.000	-2.814	40	.008	-.30	.11	-.52	-8.57E-02
	Equal variances not assumed			-3.102	22.000	.005	-.30	9.81E-02	-.51	-.10
ConfidenceQ12	Equal variances assumed	16.351	.000	-2.924	40	.006	-.2105	7.201E-02	-.3561	-6.50E-02
	Equal variances not assumed			-3.169	26.990	.004	-.2105	6.644E-02	-.3469	-7.42E-02
UserTimeQ12	Equal variances assumed	9.050	.005	4.320	40	.000	40.54	9.38	21.57	59.50
	Equal variances not assumed			4.662	28.256	.000	40.54	8.70	22.73	58.34

Table 54 T-Test – By Question - Washout

APPENDIX M: ETHICAL RESEARCH

The author is aware of the possible ethical anomaly in many scientific studies concerning possible consequences of research. While it is important to consider the ethical considerations of research on the subjects, the standard "informed consent procedures" are completely uncontroversial for an experiment that has no chance of harming anyone. What may be controversial is the possible consequences of the research if a foreseeable outcome may make it easier to "make people suffer."

Professor Nagy, a member of the Supervisory Committee, posed the following questions: If any country can wage war with near-zero casualties, are they more likely to do so? If research makes it more efficient, does it raise the prospect that more people will suffer? People likely to suffer include friendly and enemy soldiers and non-combatants. As illustrated by the following words that Professor Nagy submitted for this appendix, he disputes the analysis presented.

> *I commend the author for devoting an appendix to ethics in his dissertation, but must dissent strongly from the facts and conclusions put forward by the author's expert interdisciplinary panel. My rebuttal and call for a discussion with the interdisciplinary panel is in preparation for a submission to Joint Services Conference on Professional Ethics (JSCOPE). My dissent is at www.gwu.edu/~nagy/jscope2001.htm - Nagy [NAG01]; [NAG99].*

M.1 Historical Consequences

Mike Colarusso [COL00], a history professor at the United States Military Academy, offers the following perspective paraphrased below:

There is not a consistent correlation between the supposed "ease" technology brings to force projection and a proclivity to use it. For example, the United States response to the destruction of two American embassies cannot be characterized as anything but a measured one. Additionally, examine the supposed "ease" of such a response. It requires the repositioning of American air or naval assets, the deployment of our young people far from home, and carries with it the possibility that the enemy may respond with further

attacks upon Americans or American interests. This hardly qualifies as the exercise of force because it is easy and runs no risk of consequences.

Secondly, would the response without cruise missile technology really have been a battalion of Rangers? More likely, it would have been 16-inch shells from a battleship or conventional bombs (Libya in 1986). Technology does not explain why the United States responded, but it certainly made it possible to respond in a more measured way. In this case technology prevented a thoughtless application of brute force rather than encouraging it.

Finally, consider the way the United States has applied force and continues to apply it. It is really predicated much less upon ease rather than upon an assessment (admittedly unsound at times) of what will be gained. Some administrations do that well; others do it poorly. Examine our response to Serbia's Kosovo policy in 1999. The administration did not want to use ground power and would not have used ground power if air and cruise missile attacks did not trigger a Serbian withdrawal. In this instance, the supposed "ease" with which devastating force could be brought to bear possibly contributed to the use of force. Overall, this highlights the impact of poor policy and planning rather than technological ease. If the Kosovo crisis had occurred earlier in terms of technology, the response might have been much the same, but with crude B-52 attacks rather than precise F-117s and cruise missiles. The precision of technologically enhanced weapons may actually reduce suffering.

The ultimate example is that ICBMs were never used by the United States. Such use may be the decisive ease in just pushing a button. A counter argument may be that it was a bipolar world and the U.S. could not use ICBMs without eliciting a devastating Soviet response. Again, having assessed the consequences of using the technology, the U.S. would never determine that it was in its best interests to use ICBMs in a first strike role. Ultimately, whether overwhelming destructive force is used is not so much a function of the means with which to apply it but the consequences of applying it. That has always been the case, in both a bi-polar world and in one where the U.S. is the "sole" superpower.

Simply because the most horrible consequence (nuclear destruction) has been removed does not mean that there are not consequences to our actions.

In particular, capitalist democracies are always most aware of the consequences their actions have upon domestic and world opinion. Does the United States possess the resources to go in to Afghanistan and "get" Osama Bin Laden? Yes. Can the US eliminate Castro reasonably quickly through the use of high technology systems? Yes. Can the U.S. afford the consequences of such acts upon its relations with other, albeit non-superpower, states? No. Would the U.S. public accept such behavior? Again, no.

M.2 Moral Consequences

Father Jerry Deponai [DEP00], a chaplain at West Point offers this response.

> *It is fairly obvious that a person is morally responsible for their acts. If the initial intent and content of a particular research project by a particular person is within moral parameters, the initialing person is not morally responsible for the subsequent "redirection", "moral misapplication" of that research/technology. The person who "redirected" or "morally misapplied" the technology/research is morally responsible.*

Finally, Gary Solis [SOL00], a law professor at West Point offers these opinions, paraphrased below:

There several techniques by which legitimate and ethical scholarship may be warped, even perverted, in application. This is the case particularly when that scholarship is employed beyond the ambit of its original intent. When great political or military power is in possession of potentially destructive scholarship - nuclear energy comes to mind - safeguards for the ethical and constructive use of such scholarship are particularly called for.

Nevertheless, such safeguards have usually proven far from foolproof. Indeed, on occasion we have seen purposeful circumvention of safeguards not by criminals and enemies, but by the very agencies entrusted with the safeguards for proper use.

What is the conscientious researcher, scientist, or scholar to do? The conclusion may well be that there is no guaranteed prophylactic measure that can be uniformly effective. The best one may strive for in an imperfect world is honest and ethical work applied in the honorable ways envisioned by the scholar. If future mischief ensues, safeguards having been purposefully circumvented, no measure will have been effective. One can hope for the application and utilization of scholarship in accord with the high philosophic ideals of the state making use of that scholarship. Nevertheless, there can be no guarantees.

M.3 Summary

In short, national policy and the consequences of public and international opinion have greater impact on decisions to use force than the ease with which it may be applied. Technological advances such as precision weapons probably reduce suffering rather than increase it. Finally, the scientist who conducts research within moral parameters cannot prevent the redirection of the work and no prophylactic measure can prevent future mischief.