EVALUATION OF THE PuSHMe REGIONAL MAYDAY SYSTEM OPERATIONAL TEST

NOTE TO READER:

THIS IS A LARGE DOCUMENT

Due to its large size, this document has been segmented into multiple files. All files separate from this main document file are accessible from links (blue type) in the table of contents or the body of the document.

EVALUATION OF THE PuSHMe REGIONAL MAYDAY SYSTEM OPERATIONAL TEST

WA-RD 443.1

Final Report
September 1997

Washington State
Department of Transportation

Washington State Transportation Commission
Planning and Programming Service Center
in cooperation with the US. Department of Transportation
Federal Highway Administration

Final Report
Research Project T9903,60 Task 44
PuSHMe-Emergency Response Operational Test

Evaluation of the PuSHMe Regional Mayday System Operational Test

by

Mark Haselkorn
Principal Investigator

Jan Spyridakis	David Gautschi	Kathleen Semple	Caitlin O'Connor
Investigator,	Investigator,	Research Consultant,.	Research Assistant,
Usability	Marketability	Performance and Usability	Market and Institutional

Washington State Transportation Center (TRAC)
University of Washington, Box 354802
University District Building
1107 NE 45th Street, Suite 535
Seattle, Washington 98 105-463 1

Washington State Department of Transportation
Technical Monitor
Morgan Balogh
Intelligent Transportation Systems (ITS) Engineer
Advanced Technology Branch

Prepared for

Washington State Transportation Commission
Department of Transportation
and in cooperation with
U.S. Department of Transportation
Federal Highway Administration

September 1997

DISCLAIMER

Table of Contents

Executive Summary 1

Chapter 1

Introduction 6

1.1 Purpose 6

1.2 Background 6

1.3 PuSHMe System Descriptions 6
- 1.3.1 Overview 6
- 1.3.2 Motorola, Inc. System Description 7
- 1.3.3 XYPOINT System Description 8

1.4 Limitations of Evaluation 9
- 1.4.1 Use of Simulation 9
- 1.4.2 Selection of Test Participants 9
- 1.4.3 Incomplete Development Stage of Test Devices 9
- 1.4.4 CSC Operator Training and Experience 10
- 1.4.5 Maintenance and Monitoring of Systems 10
- 1.4.6 Impacts on Congestion and Air Quality 11

1.5 Overview of Other In-Vehicle Emergency Response Systems Utilizing GPS Technology 11
- 1.5.1 Project Northstar 11
- 1.5.2 Colorado Mayday System 12
- 1.5.3 Ford Motor Company 12
- 1.5.4 General Motors Corporation 13

1.6 Overview of Additional Methods of Determining Vehicle Location 13

1.7 Overview of this Report 14

Chapter 2. Performance Analysis 15

2.1 Introduction 15
- 2.1.1 Overview of Tests Conducted 15
- 2.1.2 Data Sources 17
- 2.1.3 Limitations 17

2.2 User Group Deployment Test 17
- 2.2.1 Overview of the User Group Deployment Test 17

2.2.2 Objectives of User Group Deployment Test 18
2.2.3 Test Methodology 18
2.2.4 Data Analysis 19
2.2.4.1 Data Analysis of Trial Success 19
2.2.4.2 Data Analysis to Determine the Time it took for a Trial to be Completed 20
2.2.4.3 Additional Data Analysis to Determine Relationships Between Trial Success and Month of Testing Button Pushed, Weather, Location, Time of Day, and Device 21
2.2.5 Results of Motorola's User Group Deployment 2l
2.2.5.1 Motorola's Overall Success of Trials 22
2.2.5.2 Motorola Time to Complete Trial 22
2.2.5.2.1 User Reported Time from Button Push to Call Acknowledgment 23
2.2.5.2.2 Computer Generated Data - Time from Call Receipt at CSC to Call Verification 23
2.2.5.3 Motorola-Relationship between Trial Success and Month of Testing 25
2.2.5.4 Motorola-Relationship between Trial Success and Time of Day 26
2.2.5.5 Motorola-Relationship between Trial Success and Button Pushed 28
2.2.5.6 Motorola-Relationship between Trial Success and User Reported Weather Conditions 28
2.2.5.7 Motorola-Relationship between Trial Success and User Reported Trial Location 30
2.2.5.8 Motorola-Relationship Between Trial Success and Individual In-Vehicle Unit 31
2.2.6 Results of XYPOINT's User Group Deployment 33
2.2.6.1 XYPOINT's Overall Success of Trials 33
2.2.6.2 XYPOINT User Group Deployment Results-Time to Complete Trial 33
2.2.6.2.1 User Reported Time from Button Push to Trial Completion 34
2.2.6.2.2 Computer Generated Data-Time from Call Receipt at CSC to Call Verification 35
2.2.6.3 XYPOINT-Relationship between Trial Success & Month of Testing 36
2.2.6.4 XYPOINT-Relationship between Trial Success and Time of Day 38
2.2.6.5 XYPOINT - Relationship between Trial Success and Button Pushed 39
2.2.6.6 XYPOINT - Relationship between Trial Success, and User Reported Weather Conditions 41
2.2.6.7 XYPOINT - Relationship between Trial Success and User Reported Trial Location 42

2.2.6.8 XYPOINT - Relationship between Trial Success and Individual In-Vehicle Unit 44
2.2.7 Conclusions and Discussion 45
2.2.7.1 Success Rate of Systems 45
2.2.7.2 Response Time 46
2.2.7.3 Other Conclusions 47

2.3 Simulated Service Delivery Test 49
2.3.1 Objectives of Test 49
2.3.2 Test Methodology 49
2.3.3 Data Analysis 50
2.3.4 Results: Time to Dispatch Service 52
2.3.4.1 Motorola Time to Dispatch Service 52
2.3.4.2 XYPOINT Time to Service Dispatch 53
2.3.5 Conclusions: Time to Service Dispatch 54
2.3.6 Results: Quality of Information 55
2.3.6.1 Dispatcher Results 56
2.3.6.1.1 AAA Dispatcher Results 56
2.3.6.1.2 Ring County Police Dispatcher Results 60
2.3.6.2 Service Provider Results 63
2.3.6.2.1 Motorola Service Provider Quality of Information Results 64
2.3.6.2.2 XYPOINT Service Provider Quality of Information Results 68
2.3.7 Conclusions: Quality of Information 72

2.4 Specific Test 1: Dropped Carrier Test 73
2.4.1 Objective of Test 73
2.4.2 Test Methodology 73
2.4.3 Data Analysis 74
2.4.4 Results 74
2.4.4.1 Controlled Dropped Carrier Testing Results 74
2.4.4.2 Applicable Results from User Croup Deployment 75
2.4.5 Discussion and Conclusions 78

2.5 Specific Test 2: Topographic Interference Test 78
2.5.1 Objective 79
2.5.2 Test Methodology 79
2.5.3 Data Analysis 79
2.5.4 Results 80
2.5.5 Results by Type of Barrier 80
2.5.5.1 Trials Conducted in Parking Garages 80
2.5.5.2 Trials Conducted in "Urban Canyons" 81
2.5.5.3 Trials Conducted in Forests 81
2.5.5.4 Trials Conducted in Open Terrain 81
2.5.6 Discussion and Conclusions 81

2.6 Specific Test 3: Moving Vehicle Test 82
2.6.1 Objective 82
2.6.2 Test Methodology and Data Analysis 82
2.6.2.1 Test Methodology and Data Analysis for XYPOINT 82
2.6.2.2 Test Methodology and Data Analysis for Motorola 83
2.6.3 Results 83
2.6.3.1 XYPOINT Moving Vehicle Test Results 83
2.6.3.2 Motorola Moving Vehicle Test Results 84
2.6.4 Discussion and Conclusions 84

2.7 Specific Test 4: Location Specific Test 84
2.7.1 Objective and Background 85
2.7.2 Methodology 85
2.7.3 Data Analysis 86
2.7.4 Results 86
2.7.5 Discussion and Conclusions 89

2.8 Specific Test 5: Remote CSC Operator Test 90
2.8.1 Objective 90
2.8.2 Test Methodology 90
2.8.3 Data Analysis 90
2.8.4 Results 91
2.8.5 Discussion and Conclusions 91

Chapter 3. Usability Analysis 93

3.1 Overview of Usability Analysis 93

3.2 Methods 93

3.3 Demographic Characteristics of Respondents 93
3.3.1 Motorola Respondents 94
3.3.2 XYPOINT Respondents 96

3.4 Data Analysis 98

3.5 Results99
3.5.1 Motorola Results 99
3.5.1.1 Motorola User Perceptions of Ease of Use: General Ease of Use 99
3.5.1.2 Motorola User Perceptions of Ease of Use: Microphone Versus Handset 100
3.5.1.3 Motorola User Perceptions of Security/Safety 103
3.5.1.4 Motorola User Perceptions of Reliability & Consistency 104

3.5.1.5 Additional Motorola User Perceptions of the Testing or System 106
3.5.1.6 Relationships between Motorola Demographic Characteristics and Perceptions of the System 107
3.5.1.6.1 Motorola Significant Differences in Responses by Gender, Age, Income, and Occupation 107
3.5.1.6.2 Motorola Significant Differences in Responses by Cellular Phone Ownership, Use of Pagers, & Home Security System Subscription 108
3.5.2 XYPOINT Results 108
3.5.2.1 XYPOINT User Perceptions of Ease of Use 108
3.5.2.2 XYPOINT User Perceptions of Security/Safety 111
3.5.2.3 XYPOINT User Perceptions of Reliability & Consistency....... 112
3.5.2.4 Additional XYPOINT User Perceptions of the Testing or System 114
3.5.2.5 Relationships between XYPOINT Demographic Characteristics and Perceptions of the System 114
3.5.2.5.1 XYPOINT Significant Differences in Responses by Gender, Age, Income, and Occupation 115
3.5.2.5.2 XYPOINT Significant Differences in Responses by Cellular Phone Ownership, Use of Pagers, & Home Security Service Subscription 115

3.6 Discussion and Conclusions 116
3.6.1 Discussion and Conclusions for the Motorola System 116
3.6.2 Discussion and Conclusions for the XYPOINT System 117

Chapter 4. Market Analysis 118

4.1 Objectives of Market Analysis 118

4.2 Study Description 118

4.3 General Overview of the Study Methodology 118

4.4 Methodology 120
4.4.1 Stage 1 Methodology 120
4.4.2 Stage 2 Methodology 121
4.4.3 Stage 3 Methodology 122

4.5 Data Analysis 122
4.5.1 Stage 1 Data Analysis 122
4.5.2 Stage 2 Data Analysis 122
4.5.2.1 Price Effects 123
4.5.2.2 Grouping within the Sample 123

4.5.2.3 Estimation of Preferences 124
4.5.2.4 Procedure 124
4.5.3 Stage 3 : Surveys of Non Field Test Participants 126

4.6 Results126
4.6.1 Stage 1 Results 126
4.6.2 Stage 2 Results 128
4.6.2.1 Aggregate 129
4.6.2.2 Participants vs. Non-Participants............................ 129
4.6.2.3 Participant Sub Groups 130
4.6.2.4 Non-Participant Sub Groups 133
4.6.3 Stage 3 Results 134

4.7 Conclusions 137

Chapter 5. Institutional Issues Analysis............................ 139

5.1 Introduction 139

5.2 Scope............................ 139

5.3 Background 139

5.4 Methodology............................ 140

5.5 Results141
5.5.1 Literature Review............................ 141
5.5.1.1 E-911 and Cellular Calls............................ 141
5.5.1.2 Other Benefits and Issues............................ 143
5.5.2 Focus Groups 143
5.5.3 Partner Interviews 145
5.5.3.1 Institutional Issues Encountered in the Operational Test 145
5.5.3.2 Institutional Issues Affecting Future Deployment.................... 146

5.6 Discussion and Conclusions............................ 147
5.6.1 Standalone Emergency Services as a Public Responsibility.................. 148
5.6.2 Standalone Emergency Services as a Private Opportunity...................... 149
5.6.3 Including Related Non-Emergency In-Vehicle Services Under the Public Mandate............................ 149
5.6.4 Packaging Emergency and Non-Emergency In-Vehicle Services as a Private Service............................ 150
5.6.5 Coordinating Public and Private Sector Partners in the Delivery of Emergency Services............................ 151

List of Figures and Tables

FIGURES

Figure 2.1 - Overview of Steps in Simulated Service Delivery Test............................ 16
Figure 2.2 - Motorola User Group Deployment: Successful and Unsuccessful Trials................ 22
Figure 2.3 - Motorola User-Reported Time to Operator Acknowledgement.............................. 23
Figure 2.4 - Motorola CSC Operator Time from Call Reception to Verification....................... 24
Figure 2.5 - Motorola Trial Success by Month (Number)............................ 25
Figure 2.6 - Motorola Trial Success by Month (%)............................ 26
Figure 2.7 - Motorola Trial Success by Time of Day (Number)............................ 27
Figure 2.8 - Motorola Trial Success by Time of Day (%)............................ 27
Figure 2.9 - Motorola Trial Success by Button Pushed............................ 28
Figure 2.10 - Motorola Trial Success by Weather (Number)............................ 29
Figure 2.11 - Motorola Trial Success by Weather (%)............................ 29
Figure 2.12 - Motorola Trial Success by Location (Number)............................ 30
Figure 2.13 - Motorola Trial Success by Location (%)............................ 31
Figure 2.14 - Motorola Trial Success by Unit (Number)............................ 32
Figure 2.15 - Motorola Trial Success by Unit (%)............................ 32
Figure 2.16 - XYPOINT Overall Success of Trials............................ 33
Figure 2.17 - XYPOINT User-Reported Time from Button Push to Trial End.......................... 34
Figure 2.18 - XYPOINT CSC Operator Time from Call Reception to Verification.................. 35
Figure 2.19 - XYPOINT Trial Success by Month (Number)............................ 36
Figure 2.20 - XYPOINT Trial Success by Month (%)............................ 37
Figure 2.21 - XYPOINT Trial Success by Time of Day (Number)............................ 38
Figure 2.22 - XYPOINT Trial Success by Time of Day (%)............................ 39
Figure 2.23 - XYPOINT Trial Success by Initial Button Pushed............................ 40
Figure 2.24 - XYPOINT Trial Success by Confirmation Button Pushed.............................. 40
Figure 2.25 - XYPOINT Trial Success by Weather (Number)............................ 41
Figure 2.26 - XYPOINT Trial Success by Weather (%)............................ 42
Figure 2.27 - XYPOINT Trial Success by Location (Number)............................ 43
Figure 2.28 - XYPOINT Trial Success by Location (%)............................ 43
Figure 2.29 - XYPOINT Successful/Unsuccessful Trials by Unit (Number)............................ 44
Figure 2.30 - XYPOINT Successful/Unsuccessful Trials by Unit (%)............................... 45
Figure 2.31 - AAA Dispatcher Attitudes Towards Usefulness of Incident Information............. 56
Figure 2.32 - AAA Dispatcher Attitudes Towards Accuracy of Location Information............... 57
Figure 2.33 - AAA Dispatcher Attitudes Towards PuSHMe Incident Information Compared to Incident Information Received from Cellular 911 Calls............................ 58
Figure 2.34 - AAA Dispatcher Attitudes Towards PuSHMe Location Information Compared to Location Information Received from Cellular 911 Calls............................ 59
Figure 2.35 - King County Police Dispatcher Attitudes Towards Usefulness of PuSHMe Incident Information 60
Figure 2.36 - Ring County Police Dispatcher Attitudes Towards Accuracy of PuSHMe Location Information............................ 61

Figure 2.37 - King County Police Dispatcher Attitudes Towards PuSHMe Incident Information Compared to Incident Information Received from Cellular 911 Calls 62
Figure 2.38 - King County Police Dispatcher Attitudes Towards PuSHMe Location Information Compared to Location Information Received from Cellular 911 Calls 63
Figure 2.39 - Simulated Service Provider Attitudes Towards the Usefulness of Motorola Incident Information 64
Figure 2.40 - Simulated Service Provider Attitudes Towards the Accuracy of Motorola Location Information 65
Figure 2.41 - Simulated Service Provider Attitudes Towards the Motorola Incident Information System Compared to Incident Information Received from Cellular 911 Calls 66
Figure 2.42 - Simulated Service Provider Attitudes Towards Motorola's Location Information Compared to Information Received from Cellular 911 Calls 67
Figure 2.43 - Simulated Service Provider Attitudes Towards the Usefulness of XYPOINT Incident Information 68
Figure 2.44 - Simulated Service Provider Attitudes Towards the Accuracy of XYPOINT Location Information 69
Figure 2.45 - Simulated Service Provider Attitudes Towards XYPOINT's Incident Information Compared to Incident Information Received from Cellular 911 Calls 70
Figure 2.46 - Simulated Service Provider Attitudes Towards XYPOINT's Location Information Compared to Information Received from Cellular 911 Calls 71
Figure 2.47 - User Croup Deployment Trials with Lost Connections 75
Figure 2.48 - Trials With Single or Multiple Call IDs - Number Losing Contact During Trial... 76
Figure 2.49 - Trials With Single or Multiple Call IDs - Percent Losing Contact During Trial... 77
Figure 2.50 - Location Specific Test Results - Motorola Histogram 87
Figure 2.5 1 - Location Specific Test Results - XYPOINT Histogram 87
Figure 2.52 - Location Specific Test Results - Scatter Diagram of 30 Meters Away from Monument 88
Figure 3.1 - Motorola Usability Respondents By Education 95
Figure 3.2 - Motorola Usability Respondents By Income 95
Figure 3.3 - Motorola Usability Respondents By Occupation 96
Figure 3.4 - XYPOINT Usability Respondents By Education 97
Figure 3.5 - XYPOINT Usability Respondents By Income 97
Figure 3.6 - XYPOINT Usability Respondents By Type of Occupation 98
Figure 4.1 - Stage 3 Survey Results, Previous Knowledge of IVERS 135
Figure 4.2 - Stage 3 Survey Results, Perceived Need for IVERS 136
Figure 4.3 - Stage 3 Survey Results, Desire for Regulation of IVERS 136

TABLES

Table 2.1 - Motorola Trial Ratings for User Croup Deployment 20
Table 2.2 - XYPOINT Trial Ratings for User Croup Deployment 20
Table 2.3 - Motorola Simulated Service Delivery Test Results 52
Table 2.4 - XYPOINT Simulated Service Delivery Test Results 54
Table 2.5 - Results of Dropped Carrier Test (Motorola/RSPI I System Only) 74
Table 2.6 - Results of Topographic Interference Test (XYPOINT) 80

Table 2.7 - Results of Topographic Interference Test (Motorola)............................... 80
Table 2.8 - Moving Vehicle Tests – XYPOINT.............................. 83
Table 2.9 - Moving Vehicle Tests – Motorola............................... 84
Table 2.10 - Results of Remote CSC Operator Test............................... 91
Table 3.1 - Motorola Perceptions of General Ease of User of the System................................ .l00
Table 3.2 - Motorola User Perceptions of Ease of Use of Handset Versus Microphone............ 101
Table 3.3 - Motorola User Preferences for Handset Versus Microphone................................ .. 102
Table 3.4 - Motorola User Perceptions of Safety and Security Benefits of System.................. 103
Table 3.5 - Motorola User Perceptions of Reliability.............................. 105
Table 3.6 - Motorola User Perceptions of Consistency.............................. 105
Table 3.7 - Additional Motorola User Perceptions of the Testing or System............................ 106
Table 3.8 - XYPOINT User Perceptions of Ease of Use............................... .... l10
Table 3.9 - XYPOINT User Perceptions of Security/Safety.............................. .... l11
Table 3.10 - XYPOINT User Perceptions of Reliability and Consistency............................... 113

List of Appendices

APPENDIX A:	GENERAL SEQUENCE OF EVENTS FOR PERFORMANCE ANALYSIS TESTS
APPENDIX B:	DATA COLLECTED IN PERFORMANCE ANALYSIS TESTS
APPENDIX C:	USER RESPONSE FORMS, MOTOROLA AND XYPOINT
APPENDIX D:	SIMULATED SERVICE DELIVERY INSTRUCTIONS PROVIDED TO USERS AND OPERATORS
APPENDIX E:	CONVERTING LOCATION DATA FOR LOCATION SPECIFIC TEST
APPENDIX F:	REMOTE CSC TRIAL LOGS
APPENDIX G:	USABILITY SURVEYS
APPENDIX H:	MOTOROLA USER GROUP DEPLOYMENT - BUTTON PUSH BY SCENARIO
APPENDIX I:	XYPOINT USER GROUP DEPLOYMENT - BUTTON PUSH BY SCENARIO
APPENDIX J:	USABILITY CORRELATIONS
APPENDIX K:	USABILITY DEMOGRAPHICS
APPENDIX L:	STAGE 1 MARKET SURVEY
APPENDIX M:	STAGE 2 DESIGN MATRIX AND SAMPLE CARDS
APPENDIX N:	STAGE 3 MARKET SURVEY
APPENDIX O:	CONSISTENCY SCORES

APPENDIX P: STANDARD DEVIATIONS OF TEST CARD RANKINGS FROM MODEL PREDICTIONS

APPENDIX Q: AVERAGE BACKGROUND CHARACTERISTICS

APPENDIX R: FREQUENCY PROFILES RANKED 1,2,3

APPENDIX S: LUTE REPORT

APPENDIX T: LETTER TO PARTNERS REGARDING INSTITUTIONAL ISSUES IMPACTING THE OPERATIONAL TEST

APPENDIX U: FCC REPORT AND ORDER AND FURTHER NOTICE OF PROPOSED RULEMAKING

EXECUTIVE SUMMARY

This executive summary presents the highlights of an independent evaluation of the Puget Sound Emergency Response Operational Test (PuSHMe), a test of regional in-vehicle mayday systems. The PuSHMe project was a joint effort by government, private industry and academia to implement and test two systems, one from Motorola and the other from XYPOINT, that allowed test groups of drivers to signal a need for in-vehicle emergency assistance to a monitored response center. Evaluation was performed in four areas: (1) system performance, (2) usability, (3) market analysis and (4) institutional issues.

SYSTEM PERFORMANCE

There were three types of performance tests: (1) User Group Deployment, (2) Simulated Service Delivery, and (3) Specific Feature Tests.

USER GROUP DEPLOYMENT

The user group deployment test included over 5,000 trials conducted between November 1995 and May 1996 that produced analyzable data; nearly two thirds of these were trials of the XYPOINT system. There was a 70.5% success rate of all trials (88% successful or probably successful for Motorola, 66% successful or probably successful for XYPOlNT). Given various test limitations and the state of infrastructure and operations deployment, it was concluded that these systems could approach 100% successful operation in a true market deployment.

In terms of "response time" (time to answer and exchange basic information), the vast majority of calls were handled within two minutes and almost no calls took longer than five minutes. These rapid response times demonstrated a potential to facilitate emergency response to on-road incidents.

SIMULATED SERVICE DELIVERY

The Simulated Service Delivery Test incorporated simulated dispatch and delivery of emergency services. This test allowed us to evaluate (1) time to dispatch and (2) quality of information.

For time to dispatch, it was not surprising that it took about twice as long, approximately twelve minutes, for a text-based system (XYPOINT) to exchange pre-dispatch information between volunteer and response center operator than a voice-based system (Motorola). For both systems, the information exchange between user and response center operator (approximately four minutes for voice, nine minutes for text) did not significantly shorten the time it took the response center operator to communicate the incident information to service providers as compared to the lime it takes a user to communicate this information directly. In other words, not surprisingly, the use of an intermediary operator will add time to the transaction. Of course time is not the only issue. The pre-screening of calls has other benefits that may outweigh the possible time lost, such as the elimination of unnecessary calls to E-911 centers and better location information.

For quality of information, dispatchers and service providers found the incident information useful and the location information reasonably to very accurate. However, both the dispatchers and service providers found the quality of the incident and location information they received to be about the same as that which they currently receive from cellular 911 calls.

SPECIFIC FEATURE TESTS

A series of tests were conducted to evaluated specific aspects of the systems.

From the location specific test we learned that both PuSHMe systems produced fairly reliable, fairly accurate locations that in most cases would be helpful in support of the delivery of emergency services. However, we also learned that (1) differential correction for the Motorola system was probably not functioning for all or part of the PuSHMe test, and (2) the XYPOINT system produced a relatively large number of "outlier" locations with several trials more than 80 meters off of the known location.

From the Dropped Carrier test, we learned that: (1) the interface of the response center workstation should be reevaluated to determine if it needs to be modified to enable better handling of dropped calls, and (2) attention must be paid to operator training to ensure that operators are familiar with the procedures required to properly handle dropped calls.

From the remote operator test we learned that operators in Phoenix could handle emergency calls from Seattle drivers nearly as well as operators in Seattle.

From the moving vehicle test we learned that both PuSHMe systems were able to track moving vehicles reasonably well.

From the topographic interference test we confirmed that, as expected, both systems experienced difficulties in conditions with overhead obstructions.

USABILITY

The usability analysis of the PuSHMe systems was favorable and also produced information that can help in future design modifications. Usability testing was performed on the devices as they existed at the time of the operational test. The sample population was drawn from participants in the user group deployment which consisted of people with high levels of education, income, and technological sophistication.

Response to Motorola's in-vehicle emergency response system was very favorable with most users finding the device easy to use, reliable, and consistent. Use of the handset appeared to lead to clearer communication (particularly with the operator being better able to hear the user), though most respondents reported that they could hear the operator and vice versa using either the handset or microphone. In fact, most respondents preferred to use the microphone, which enabled hands-free communication. Most respondents also expressed a very favorable attitude toward the safety and/or security that this system could offer, although several identified situations in which the system would be unable to help them. In addition, slightly more than half of respondents were surprised by a system "glitch" during operation.

Response to XYPOINT's in-vehicle emergency response was favorable in many respects and unfavorable in a few. Most respondents reported that they found the device relatively easy to use and handle, although a number did comment that the cords tangled easily and that it was awkward to set-up the device in their vehicles. In addition, many respondents commented that they sometimes missed seeing new messages appear on the screen. A majority of respondents agreed that this system would provide some feelings of safety or security to them, though this judgment was tempered by the fact that most users found that the system performed inconsistently. The most common issues brought up by respondents were that the system was not consistent in how it worked and that they would have liked the system to have allowed them to speak with the operator.

MARKET ANALYSIS

The market analysis evaluated market demand for in-vehicle emergency response systems (IVERS) such as those used in the PuSHMe operational test. Our goals were to: (1) provide guidance to the PuSHMe partners as they refine the designs of the systems that they will ultimately bring to market, and (2) provide information to the government and other interested parties on likely market scenarios for these products.

We discovered that:

- Cost is the key factor in the marketability of an IVERS, with purchase cost having far more impact than usage cost. This implies a clear pricing strategy favoring usage fees for most market segments. This is not to say that people do not consider function--they do, but they do not generally want to trade higher cost for additional functionality.

- Use of an IVERS fostered a sensitivity to the value of additional information (e.g. directions), while non-users were more sensitive to service features (e.g. towing). This implies that initial marketing to encourage adoption of these systems might emphasize bundled related services, while later efforts to maintain market share might focus on enhanced information services.

- Influences of purchase cost, usage fees, product design, information, and service levels are not at all consistent across the subgroups. This indicates the critical importance to producers of IVERS of carefully segmenting the markets that they intend to supply.

- Skepticism exists about the economic value of IVERS. Coupled with the consistently revealed sensitivity to purchase cost , suppliers of IVERS would be most successful if they could bundle the IVERS into the price of other purchases -- most likely that of an automobile.

INSTITUTIONAL ISSUES

The primary institutional issues surrounding deployment of an IVERS involve the public/private partnership which must be evolved to make such a system feasible. These issues extend beyond protocols and guidelines which can be developed to define agreed upon interaction between the public and private sector entities. At the highest level, this partnership hinges upon assumptions about the role and cost of government and the benefits of citizenship. Like health care and other governmental activities that are viewed as public goods, transportation is grappling with the complex question of what level of service should be provided equally to all citizens and which services beyond that level should be provided non-uniformly on other bases such as ability/willingness to pay or to form a car pool.

We concluded that:

- Taken in isolation, emergency response to automobile accidents will likely continue to be seen as falling within the spectrum of transportation services that are public responsibility.

- Some high end market for standalone private emergency response services can likely be generated through appeals to higher levels of personal security, but as a standalone service, emergency response is probably insufficient to justify the technology and infrastructure expense of a private service.

- In a climate of reduced government service and overworked E-91 1 centers, it seems highly unlikely that enhanced non-emergency in-vehicle services such as helping drivers when they are lost will be seen as falling within the spectrum of public responsibility.

- Packaged emergency and non-emergency in-vehicle services represent a clear commercial opportunity, in fact, there may be sufficient value in the non-emergency services to make them a viable product independent of emergency services

- There are two somewhat conflicting visions concerning commercial delivery of packaged emergency and non-emergency in-vehicle services:
  (1) a vision that focuses on non-emergency services and downplays the coordination issue. The private response center simply passes through calls that fit an agreed upon definition of an emergency. The advantages of this vision are the simplification of institutional implementation issues and reduced time for emergency calls to reach the public dispatcher.
  (2) a vision where the private response center is a close partner with the E-911 centers. The private center provides service to the E-911 centers by pre-screening calls and providing additional valuable information. In return, the E-911 centers are willing to work out mutually beneficial protocols and standards for the private to public hand off. The advantages of this vision are the sharing of valuable information and resources, as well as the appearance to the user of a single service provider.

- There are serious obstacles to the collaborative vision. Some of these obstacles relate to developing common communication standards and operational protocols, such as adopting a common language and schema in an environment where PSAPS themselves vary greatly. Other obstacles relate to the reluctance among public agencies to add any non-public entity into their organizational procedures since it complicates their already difficult and sensitive efforts. For example, public dispatchers do not want private operators to make any decisions that can impact the safety of public emergency response personnel.

- There is a third vision where private response groups ‘become” the public emergency service provider through outsourcing.

CHAPTER 1. INTRODUCTION

1.1 PURPOSE

This report is an independent evaluation of the Puget Sound Emergency Response Operational Test (PuSHMe), a test of regional mayday systems that allow a driver to signal his or her location and need for assistance to a response center. The project represents a partnership of both public and private organizations in&ding the Federal Highway Administration (FHWA); the Washington State Department of Transportation (WSDOT); the Washington State Patrol (WSP); David Evans and Associates, Inc. (IDEA); Motorola Inc., acting through its Space Systems Technology Group; IBI Group; XYPOINT; Response Systems Partners, Inc. (RSPI); and the University of Washington's Laboratory for Usability Testing and Evaluation (LUTE). For the purposes of this document, "evaluation team" or "evaluator" refers to the team of faculty, staff, and students at the University of Washington responsible for conducting the evaluation and writing this evaluation report (this does not include LUTE). "Partners" refers to the participants in the PuSHMe operational test.

1.2 BACKGROUND

The evaluation team worked with the PuSHMe Evaluation Working Group (EWG), a subcommittee of the partner's steering committee, to develop a detailed evaluation plan for the PuSHMe operational test. The Final Detailed Evaluation Plan was submitted on June 29, 1995 and accepted by the FHWA on November 17, 1995. During the course of the operational test, the evaluation team has worked closely with the EWG to refine this plan and data has been collected by the partners as detailed in the revised plan. In September of 1996, preliminary results were presented to the EWG for their review and comments,

1.3 PUSHME SYSTEM DESCRIPTIONS

1.3.1 OVERVIEW

The PuSHMe project was a joint effort combining the resources of government, private industry and academia to implement and test a regional mayday system. The project involved the deployment of two systems, one from Motorola and the other from XYPOINT, that allowed test groups of users to signal a need for emergency assistance to a monitored response center.

The Motorola and XYPOINT systems have both similarities and differences. Both systems use Differential Global Positioning System (DGPS) technology to locate signaling vehicles. Both systems employ a private response center to facilitate interaction between distressed vehicles and public emergency service providers. The major differences between the two systems are (1) the Motorola system provides two-way voice communication while the XYPOINT system is text based and, like a pager, employs an LCD display, and (2) Motorola relies on the standard cellular communications infrastructure while XYPOINT uses the newer Cellular Digital Packet Data system (CDPD). Motorola established its response center at the WSDOT Transportation System Management Center (TSMC) in northern Seattle, while XYPOINT established a separate response center at its offices near downtown Seattle.

1.3.2 MOTOROLA, INC. SYSTEM DESCRIPTION

Motorola, Inc.'s emergency response system combines DGPS, data communications, Analog Mobile Phone System (AMPS) cellular, mapping, and database technologies with the aim of providing automobile drivers with emergency services such as personal security (e.g. carjacking), personal emergency (e.g. medical; auto accident), roadside assistance (e.g. dead battery; breakdown), and traveler assistance (e.g. location of nearest hospital; congestion information). The system used in PuSHMe was based on the existing Motorola MotoTrack™ Emergency Response System but contained a subset of its functionality (e.g. did not include vehicle security, stolen vehicle recovery, and full travel assistance capabilities).

Participating vehicles were equipped with Motorola's in-vehicle equipment consisting of a push button device integrated with an AMPS cellular car phone. Emergency response calls were initiated by pressing one of three service request buttons (emergency, roadside assistance, or traveler assistance) located on the phone handset. The system included an emergency "panic button" that was not used in the PuSHMe test.

A receiver in the Motorola device used global positioning system (GPS) navigation messages being transmitted by the GPS satellite constellation (owned and operated by the U.S. Department of Defense and free to use by the general public) to calculate the vehicle's position and velocity. For improved position accuracy, a DGPS reference station was located at WSDOT TSMC, where PuSHMe calls from subscriber vehicles were answered at a Customer Service Center (CSC). The Telephone Company (TELCO) and AMPS cellular infrastructures provided the communications link between the CSC and the subscriber vehicles. These are existing commercial infrastructures and were not partners in the PuSHMe test, although the PuSHMe system interfaced to them.

The Motorola CSC consisted of a server and three operator workstations that allowed operators to service up to nine calls simultaneously. When the driver of a test vehicle initiated a cdl, the system automatically contacted the CSC server, established synchronization, and transmitted relevant data necessary to service the call such as position, heading, vehicle identification, and time. After receipt of a call at the **CSC,** the server placed it in the call queue. An operator answered a call by selecting it from the call queue with a mouse. The audio for the call was automatically routed to the appropriate operator headset and the vehicle's location, velocity, and heading was displayed on a map on-screen. Other pertinent data required to effectively service the call was also displayed, such as vehicle and driver attributes. The subscriber could speak with the CSC operator using either a handset or a microphone installed in the vehicle. The operator could listen in and/or talk with the vehicle to provide assistance or gather additional information. The operator also entered call and disposition notes as the call was being serviced. These notes were stored as part of the data record of the call.

1.33 XYPOINT SYSTEM DESCRIPTION

XYPOlNT's emergency response system relies on a CDPD wireless network, DGPS location technology, and an emergency response center. The emergency response center integrates multiple databases with a rules based engine to facilitate interaction between distressed vehicles and emergency response providers. The goal is to support three types of emergency calls: (1) police or highway patrol, (2) medical assistance, and (3) towing or roadside services.

For the PuSHMe test, XYPOlNT provided portable devices which allowed drivers, by pressing one of three buttons, to alert a response center operator. The buttons were labeled Police, Medical, and Auto. Upon establishing communication, the response center workstation's database was queried to yield basic data such as vehicle location, unit id, type of button pushed, and time. The XYPOINT response center operator could then contact the driver by typing or selecting predetermined messages or questions that were shown on the LCD display on the unit. The driver could answer using either a Yes or No button on the in-vehicle device.

The XYPOINT system relied on a CDPD infrastructure provided free of charge by AT&T Wireless Services. In addition, XYPOINT worked with Trimble to achieve DGPS, using

Trimble's base station in Lynnwood, Washington. Neither AT&T nor Trimble were PuSHMe partners.

1.4 LIMITATIONS OF EVALUATION

The PuSHMe evaluation was impacted by limitations imposed by the context of the operational test. During the test, devices and service response centers did not provide the full functionality and actual services for which they were designed. Because these limitations are important to understanding the scope of the evaluation, they are further described in the following subsections.

1.4.1 USE OF SIMULATION

Due to liability issues and the limited chance of an emergency occurring, there were no actual responses to real emergencies during the PuSHMe field test. During user group deployment, emergency calls were part of a series of simulated tests, and for only one group of tests did calls go beyond the response center operator. (For the simulated service delivery tests, WSDOT, the King County Police and AAA responded to simulated incidents.) Data gathered for all aspects of the project evaluation were affected by the simulated nature of the test incidents. For example, the CSC operators knew what type of tests would be occurring and approximately when. Simulated response providers also knew when the test was scheduled and that they were not responding to true emergencies. The simulated nature of the tests limited our ability to generate data that directly measured the impact of the PuSHMe systems on motorists' safety.

1.4.2 SELECTION OF TEST PARTICIPANTS

Because this operational test required a great deal of cooperation over a number of months from participants and also had to consider liability issues, subject selection was not randomized. Instead, the selection of test participants was done by the test partners and targeted cooperating organizations in the region and/or employees of the partnering organizations. While demographic factors were tracked and considered relevant, we could not select a representative sample of the Puget Sound driving public (except for parts of the market evaluation).

1.43 INCOMPLETE DEVELOPMENT STAGE OF TEST DEVICES

At the time of the PuSHMe operational test, the Motorola and XYPOINT emergency response systems were not yet fully market developed and product tested. Therefore, the evaluation results represent more an analysis of potential benefits and consumer potential than actual impact of the individual emergency response systems. In addition, this evaluation was structured to avoid directly comparing the two test emergency response systems against each other in a competitive manner. The evaluation team structured the evaluation in ways that would produce significant conclusions on the potential of these approaches to provide an important ITS service, while still considering the development stages of these test devices, systems, and infrastructures.

1.4.4 CSC OPERATOR TRAINING AND EXPERIENCE

Because the PuSHMe systems were not yet introduced into the marketplace, CSC operators were students working part-time or company staff on temporary assignment, rather than permanent employees assigned to emergency response operations. While these operators received training on CSC computer operations and data entry procedures, this training was not equivalent to that which would be received by professional emergency response operators and did not cover how to interact with customers and service providers. This was important because operators were an important source of test data, and they were called on to enter this data as they interacted with customers. In some cases, the performance measures of the system were based on the performance of the operators. For example during the simulated service delivery test, CSC operators described to the service providers details about the simulated incident and location information, and the quality of this information was evaluated. A further complication was that some specific tests lasted as few as two days, giving operators little time to assimilate test protocols. Operator inexperience, plus uncertainty over data gathering protocols, may have impacted the quality of data generated.

1.4.5 MAINTENANCE AND MONITORING OF SYSTEMS

Just as CSC operators were not trained professionals, neither were the PuSHMe mayday systems maintained at the level of fully operational products. As commercial products, the PuSHMe systems would have received a far higher degree of monitoring and maintenance than was possible during the operational test. For example, it appears that Motorola's differential correction for its GPS was not functioning properly for most or all of the life of the test. Since differential correction is well understood, this was likely a maintenance issue, and would certainly not have been left undetected in a commercially deployed product.

1.4.6 IMPACTS ON CONGESTION AND AIR QUALITY

Typically, ITS operational tests evaluate the test system's impact on congestion and air quality. This evaluation could not directly measure these impacts since no actual service was provided. It is logical to assume that reducing time to locate and reach incidents will have a positive impact on congestion and air quality, but the PuSHMe test was not designed to provide comparative data on these times.

1.5 OVERVIEW OF OTHER IN-VEHICLE EMERGENCY RESPONSE SYSTEMS UTILIZING GPS TECHNOLOGY

Although GPS technology is decades old, it has only recently been tested for use in passenger vehicles for in-vehicle emergency response or personal security. Project Northstar was a recently completed operational test of an in-vehicle traveler information and emergency notification system. Another in-vehicle response system demonstration is concurrently underway in Colorado. Recently, the Ford Motor Company began equipping some of its Lincoln Continentals with an in-vehicle emergency response system, and GM has announced the OnStar system for its 1997 front-wheel drive Cadillacs.

1.5.1 PROJECT NORTHSTAR

NYNEX Venture Company worked with other private companies, state and local agencies, and a University, to design and test an advanced traveler information, advisory, and emergency notification and security system. This project, known as Project NorthStar, was tested in the New York City metropolitan area and made use of cellular telephony, GPS technology, and a geographic and information system (GIS) (Collura, et al., 1994).

Two aspects of the project relate to in-vehicle emergency response or roadside assistance. The Emergency Roadside Service (ERS) element of the project enabled users to request emergency roadside assistance in case of vehicle breakdowns, accidents, etc. Between April and May of 1994, nine ERS calls were placed to the response center. The Personal Emergency or Position Enhanced Cellular 9 11 Services (PER) element of the project enabled equipped vehicles to dial 9 11 on their cellular phones if they were in personal emergencies. Along with the call, data providing precise location information would be transmitted to state troopers at the police dispatch centers. Between April and May of 1994, eight PER calls were placed. Perhaps because there were only a few emergency or roadside assistance calls included in this project, the success of this part of the project was not elaborated on by the project evaluators in their interim report.

1.5.2 COLORADO MAYDAY SYSTEM

Another operational test of an in-vehicle emergency response system utilizing GPS technology is currently underway in Colorado. The Colorado Mayday Emergency Response Request System utilizes the TIDGET™ Mayday System developed by NAVSYS Corporation. The TIDGET™ sensor is a low cost device which does not track GPS signals, but instead captures a brief "snapshot" of raw GPS sampled data (Brown, A. and Silva, R., 1994). The "snapshot" is transmitted to an emergency dispatch facility where it is processed to compute the location of the motorist. A map database provides support when fewer than four satellites are visible.

For this project, a full scale operational test with on-demand highway trials is being conducted over a wide area in central Colorado, covering multiple geographies and terrains. More than 2000 vehicles have been equipped with the Mayday device, with Colorado State Patrol operating a Mayday dispatch center. Many of the objectives of the test are similar to those of the PuSHMe project with regards to technical performance, institutional issues, and marketability of the Mayday system. In addition, Colorado is exploring the potential for a mayday system to be nationally implemented. Among the steps necessary to expand to national coverage are determining the structure of a nationwide public/private partnership and establishing technical standards for system components which can be utilized to expand the geographic scope of the system. The test is being run for 18 months and at its conclusion the public/private partnership will take over operations of the system.

1.5.3 FORD MOTOR COMPANY

The Ford Motor Company began offering an option on its Lincoln Continentals during the 1996 model year. Vehicles could be equipped with Lincoln's Remote Emergency Satellite Cellular Unit (Lincoln RESCU), along with a voice-activated cellular phone.

When the emergency button is pushed in a cellular phone area, the system automatically initiates a data transmission via the cellular call with the vehicle location, direction, and speed to a response center. This information will then be relayed to the local 911 system or the Lincoln Continental Roadside Assistance Program. Westinghouse Security Systems will operate the Lincoln Security Center in Irving, Texas. Westinghouse currently handles an average of 25,000 residential and business security alarm signals per day generated by systems other than RESCU (Ford Motor Company, Press Release).

Ford conducted field tests from March to May of 1995. Detailed descriptions of the field tests were not available to the PuSHMe evaluators (due to proprietary reasons), but summaries of field test results were included in a Ford Motor Company press release.

The press release indicated that 96% of test activations connected to an operator at the Westinghouse Emergency Response Center. Ford indicates that the 4% of unsuccessful activations were due to low cellular signal strength or other test limitations. Once a user pushed a button on the device, the average time it took to speak to a response center operator was 58 seconds. It took 4.4 minutes on average from dispatch of service to arrival of the chase vehicle. The average cycle time from the push of the emergency button to the chase vehicle arrival was under 11 minutes. Because detailed information concerning the test methodologies is unknown, it is difficult to compare Ford's test results to those of the PuSHMe operational test.

1.5.4 GENERAL MOTORS CORPORATION

General Motors Corporation will equip its 1997 front wheel drive Cadillacs with a device called OnStar. According to *World on Wheels* online, if an airbag deploys, an emergency signal will be sent automatically via cellular phone lines to a GM Customer Assistance Centre. If staff there are unable to communicate with the vehicle occupants over the cellular phone, emergency help will immediately be dispatched. Other planned services include: voice navigation assistance if the driver is lost or desires an alternate route; the transmission of cellular data that will unlock a vehicle at a specified time to assist a driver who has locked keys inside the vehicle; and the continuous tracking of stolen vehicles. OnStar technology will include a hands-free, voice-activated cellular service employing a hidden microphone and linked to the stereo system. The phone will recognize and respond to voice commands. The GPS antenna will be hidden in the vehicle and the OnStar transmitter/receiver will be located in the luggage compartment.

1.6 OVERVIEW OF ADDITIONAL METHODS OF DETERMINING VEHICLE LOCATION

While the PuSHMe Operational Test employed the use of GPS to determine vehicle location, several other methods of determining vehicle location are also available, including triangulation, dead reckoning, CDMA, and FM radio location.

Cellular phone triangulation uses the three closest base station sites to locate the origin of the cellular phone call by determining the latitude and longitude of the caller. In urban areas there are enough cellular base stations that several may receive the same call. If at least three base stations pick up a call, an algorithm can calculate the location based on how long the signal takes to travel to each of the three towers. According to Thomas E. Wheeler, CEO of the Cellular Telecommunications Industry Association, "between two-thirds and three-quarters of the time, the [triangulation location] system will be accurate within 25 yards." (Emmett, A,, 1996).

Dead reckoning is a technique of computing the position of a vehicle from its distance and heading measurements. It is an inertial system which uses two independent sensors from which distance and heading information can be extracted. Because each position is relative, cumulative drift errors occur on an average of 2% of the distance traveled. The system consists of a magnetic compass or gyroscope for heading and odometer for distance. Dead reckoning is often combined with map matching to correct for drift errors. It performs well when combined with highly accurate digital maps. It can also be used to back up GPS during satellite visibility outages.

The CDMA network can be used for mobile positioning. Using spread spectrum signals transmitted from each cell in the CDMA network it is possible to estimate a mobile's location. The mobile measures the arrival time differences of at least three pilot tones transmitted by three different cells. The mobile's position is then estimated by intersecting hyperbolas. The accuracy depends on the sampling rate, signal to noise ratio, and multipath amplitude.

An FM radio location system uses commercial FM radio transmitters to calculate the angles of the incoming signal and locate the cellular telephone. Since cellular phones are actually radio transmitters and receivers, all that needs to be done is to install a separate chip in the phone. This chip will receive FM radio transmissions, measure the signal, and calculate location.

1.7 OVERVIEW OF THIS REPORT

Chapters 2 and 3 of this report focus on the actual systems used in the PuSHMe Operational Test. Chapter 2 presents a performance analysis and Chapter 3 presents an analysis of the usability of these systems.

Chapters 4 and 5 of this report focus on the general concept of an in-vehicle emergency response system. Chapter 4 presents a market analysis and Chapter 5 reviews institutional issues.

CHAPTER 2. PERFORMANCE ANALYSIS

2.1 INTRODUCTION

The purpose of the performance analysis was to evaluate whether the PuSHMe mayday systems performed as designed and whether that performance was suffkient for providing effective in-vehicle mayday services. The performance tests were designed by the Evaluation Team and the Evaluation Working Croup (EWG). The tests were conducted under the direction of David Evans & Associates (DEA) All data analysis and the conclusions presented below are the work of the Evaluation Team.

2.1.1 OVERVIEW OF TESTS CONDUCTED

For the performance analysis, we looked at a number of different aspects of the two PuSHMe systems, such as how well the systems performed over a long period of time, how accurate the systems were in determining locations, how well the systems performed under challenging conditions, etc. (For an overview of the two systems, see Section 1.3 .) Because there were many different issues to evaluate, a number of different tests were conducted. There were three types of performance tests: (1) User Croup Deployment, (2) Simulated Service Delivery, and **(3)** Specific Feature Tests.

The user group deployment test generated a large amount of data about the performance of the systems over several months. This test covered the systems' operations in a variety of settings, taking a call from its initiation through the receipt of that call at the Customer Service Center (CSC). The test did not go beyond initial contact between the traveler and the CSC operator.

The simulated service delivery test involved far fewer trials than the user group deployment, but was larger in scope and came closest to testing an actual emergency response system. It involved tracking a mayday scenario from call initiation through the arrival of simulated emergency service, as shown in Figure 2.1.

The specific feature tests focused on specific performance issues related to the devices, the cellular networks, or the service centers. These tests involved fewer trials, were shorter in duration, and generally had a narrower scope. There were five specific tests: (1) Dropped Carrier, (2) Moving Vehicle, (3) Topographic Interference, (4) Location Specific, and (5) Remote Response Center.

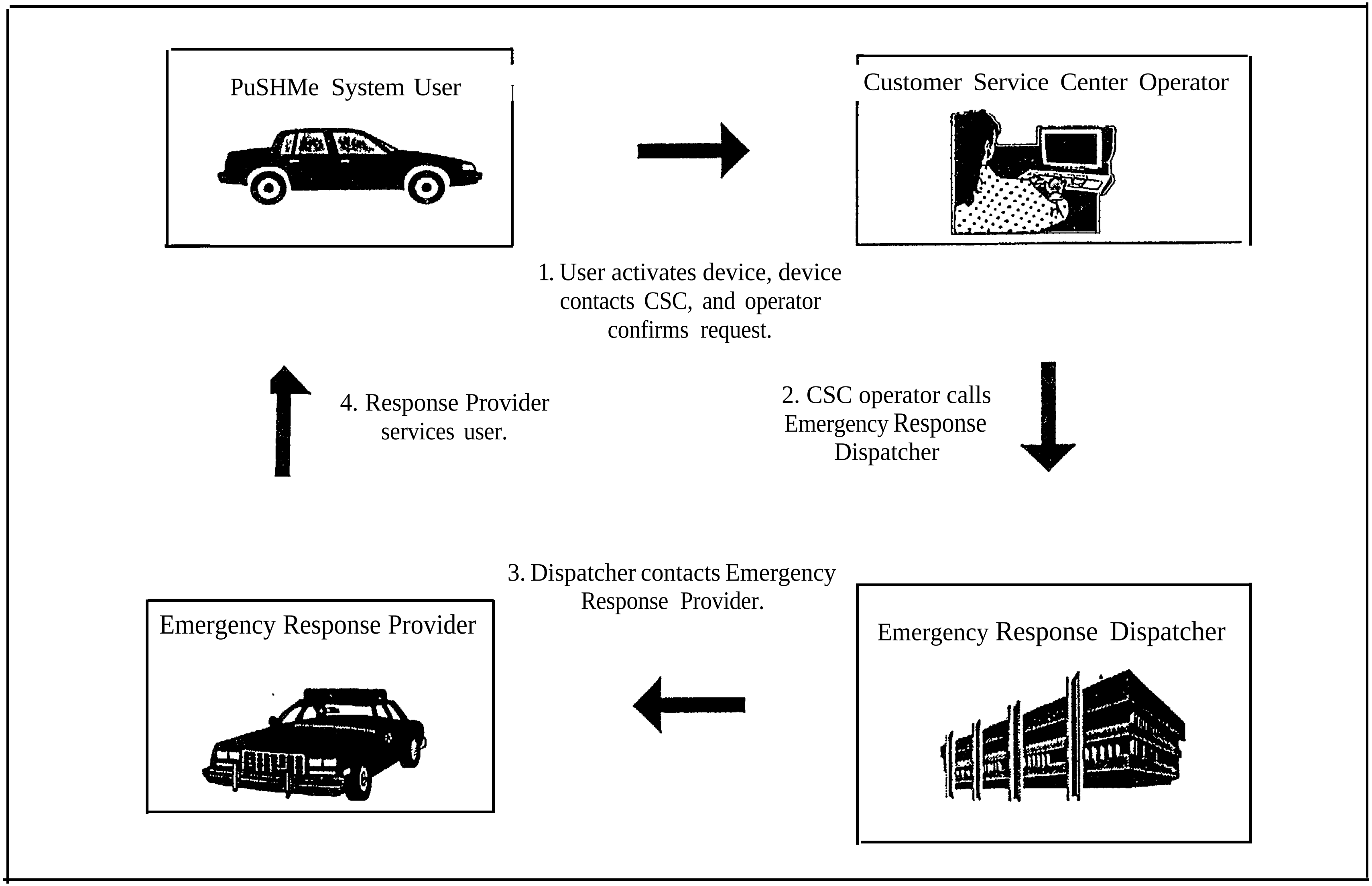


Figure 2 .1 - - Overview of Steps in Simulated Service Delivery Test.

Sections 2.2 through 2.8 present detailed discussions of each of the performance tests conducted, including objectives, methodology, analysis, results and conclusions. Appendix A presents the general sequence of events for each of the performance analysis tests.

2.1.2 DATA SOURCES

For the performance analysis tests, data was collected from four primary sources: (1) user response forms, (2) CSC operator-entered data, (3) CSC computer-generated data, and (4) simulated service provider and dispatcher logs. For the location specific test, data regarding geodetic coordinates of monuments was provided to the PuSHMe partners by the WSDOT and the city of Bellevue, Washington. The specific types of data collected by each of these sources is presented in Appendix B.

Data management, reduction, and analysis was a major effort for this study. For example, the Motorola system generated about 4 1,000 records of computer-generated data, and about 3,000 records of user response form data. The XYPOINT system generated about 30,000 records of computer-generated data, and 5,600 records of user response form data.

2.1.3 LIMITATIONS

There were a number of unavoidable limitations in the nature of the operational test and it is important to keep these in mind when interpreting the results of the PuSHMe performance evaluation. See Section 1.4 for discussion of these limitations.

2.2 USER GROUP DEPLOYMENT TEST

2.2.1 OVERVIEW OF THE USER GROUP DEPLOYMENT TEST

The user group deployment test was the largest and longest test of the PuSHMe systems. For the user group deployment test, about 190 volunteers conducted daily trials of the systems over the course of several months. The purpose of the user group deployment test was to gather large amounts of data on the performance of a significant phase of the PuSHMe systems operation--from initiation of the emergency call to verification that the call had been received and correctly understood at the Customer Service Center (CSC).

2.2.2 OBJECTIVES OF USER GROUP DEPLOYMENT TEST

The primary objectives of the user group deployment test were to evaluate: (1) the number and proportion of "successful" trials, (2) the time it took for a trial to be completed, (3) the relationship between trial success and month of testing, (4) the relationship between trial success and user-reported weather conditions, (5) the relationship between trial success and user-reported location types (urban, surburban, or rural), (6) the relationship between trial success and time of day, and (7) the trial success by individual devices. A final objective was, where possible, to compare test results to cellular 911 base data.

A secondary objective of the user group deployment was to expose .a group of volunteers to the test systems over the course of several months so that volunteers could provide feedback for the usability and market analyses (see Chapters 3 and 4).

2.2.3 TEST METHODOLOGY

For both the Motorola and XYPOINT systems, volunteers were recruited to conduct daily trials of the system. For each trial, volunteers were given a user report form to complete (see Appendix C - Sample User Group Deployment User Form). Volunteers were told to read a given scenario (e.g., "You have a flat tire and no spare.") and push one of three buttons on their in-vehicle device, according to the type of assistance they felt they would need. Volunteers were instructed to push the button during hours that the response center was operational, and they were told to push the button only one time. If a response was not made within 5 minutes, volunteers were to mark the trial as a failure on the user response form. Because the two systems are based on somewhat different technologies and approaches, the test procedures followed for the Motorola and XYPOINT systems were somewhat different, as described below.

For the Motorola system, user group deployment testing began November 15, 1995 and ended April 30, 1996, with about 40 volunteers recruited to conduct daily trials of the system. For the XYPOINT system, user group deployment testing began January 29, 1996 and ended May 3 1, 1996 with about 150 volunteers recruited to conduct daily trials of the system. With the Motorola system, users could push either an "emergency," "roadside assistance," or "traveler assistance" button. With the XYPOINT system, the users could push either a "911," "medical," or "auto" button. In each case, this button push initiated a request for assistance and the operator at the CSC then asked the user to confirm their request for assistance. With the Motorola system the operator verbally asked the user to confirm the request and then entered a verification note in the CSC data stream. With the XYPOINT system, the operator sent a text message to the user asking the user to confirm the request for assistance by pushing either the "Yes" or "No" button. Note that verification for the Motorola system was indicated by the CSC operator, while verification for the XYPOINT system was defined as the receipt of the "Yes" or "NO" message sent by the user.

2.2.4 DATA ANALYSIS

Sections 2.2.4.1 through 2.2.4.7 describe how data was analyzed to address each of the seven objectives presented in Section 2.2.2. In each case, separate analyses were performed of Motorola and XYPOINT data.

2.2.4.1 Data Analysis of Trial Success

For each trial in the User Group Deployment Test, data from the user response form and from the CSC computers was analyzed to determine the success of the trial. On the user response form, the driver indicated how the test went from his or her perspective. The CSC computers produced a different set of data. Each trial produced multiple CSC data records, and every record in a trial normally had the same call identification number. The CSC data on each trial included an indication of whether or not the call had been verified. Based on these data, each trial was assigned one of the following ratings: Successful, Probably Successful, Probably Unsuccessful, Unsuccessful, Conflicting Data 1, Conflicting Data 2, or No Trial. As mentioned earlier, the Motorola and XYPOINT systems produced somewhat different data. Tables 2.1 and 2.2 below show how data from the user response forms were linked to data from the CSCs to rate trials of each system.

The categories Conflicting Data 1, Conflicting Data 2, and No Trial were discarded from all subsequent analyses. Conflicting Data 1 was assigned when the user reported that the trial was successful but the CSC computer produced no corresponding data record (497 of 8,656 scheduled trials). Conflicting Data 2 was assigned when the user reported that the trial was unsuccessful (or never occured) but the CSC computer indicated that a trial call was received at the CSC (102 of 8,656 scheduled trials). Finally, No Trial was assigned whenever no data existed for a given trial or whenever the user indicated they never pushed an emergency call button (3,349 of 8,656 scheduled trials). This occured primarily because volunteers were unable to participate (i.e. were on vacation, had a day off, had a different car with them, etc.). Overall, 8,656 trials were scheduled with 4,708 of those scheduled trials producing usable data.

Table 2.1 Motorola Trial Ratings for User Group Deployment

	Data from Response Center		
User Report Forms	No Data for Trial	Data for Trial But Not Verified	Data for Trial and Verified
Blank/Incomplete Form	No Trial	No Trial	No Trial
All Okay	Conflicting Data 1	Probably Successful	Successful
Never Worked	Unsuccessful	Probably Unsuccessful	Conflicting Data 2
Disconnected and Reconnected	Conflicting Data 1	Probably Successful	Successful
Disconnected, Never Reconnected	Unsuccessful	Probably Unsuccessful	Probably Successful
Didn't Push Button	No Trial	Conflicting Data 2	Conflicting Data 2

Table 2.2 - XYPOINT Trial Ratings for User Group Deployment

	Data from Response Center		
User Report Forms	No Data for Trial	Data for Trial But Not Verified	Data for Trial and Verified
Blank/Incomplete Form	No Trial	No Trial	No Trial
All Okay	Conflicting Data 1	Probably Successful	Successful
Never Worked	Unsuccessful	Probably Unsuccessful	Conflicting Data 2
Pushed Button but Trial Never Completed	Unsuccessful	Probably Unsuccessful	Probably Successful
Didn't Push Button	No Trial	Conflicting Data 2	Conflicting Data 2

2.2.4.2 Data Analysis to Determine the Time it Took for a Trial to be Completed

Data was analyzed to determine how long it took for a call to be handled. Two time measurements were taken for each trial, one from the user response form and the other from the CSC computer log. On user response forms, users indicated when they pushed the button and when the call was acknowledged as received and understood by the CSC operator. From the CSC computer logs, we used the time when a call was first received and, for Motorola, the time when the call was verified by the CSC operator. For XYPOINT, we used the time that the Yes/No response to the operator's query was received. From these four times (user reported button push, user reported call answer, CSC time of call receipt, and CSC time of call verification) two time differences were measured: (1) the user reported time from button push to call answer/test completion at the CSC, and (2) the computer-generated time from call receipt at the CSC until the verification of the call. (Because the user watches were not synchronized with the CSC computers for this test, the time differences between the user reported button push and CSC verification were not measured).

2.2.4.3 Additional Data Analysis to Determine Relationships Between Trial Success and Month of Testing, Button Pushed, Weather, Location, Time of Day, and Device

Once trials were rated (see Section 2.2.4.1) as Successful, Probably Successful, Probably Unsuccessful, or Unsuccessful, tests were used to determine if there were significant differences in ratings by: (1) month of testing, (2) button pushed, (3) weather, (4) location, (5) time of day, and (5) individual device. Since the data were on ordinal scales, Kruskal-Wallis tests were used rather than ANOVAs which would have been used with interval scale data. Kruskal-Wallis analyses tested for significant differences (at an alpha-level of .05) between independent samples in trial ratings. The statistical test ensured that differences in the trial ratings were sufficiently great (when reported ***p*** values are less than or equal to .05) so as not to represent random variation.

For some of the variables identified above, we had hypotheses regarding how they might affect trial ratings. For other variables, the analyses were primarily exploratory. It was hypothesized that there might be an improvement in performance by month, since bugs in the system discovered early on could be corrected for the later months of testing. It was also hypothesized that performance might be best in urban and suburban areas, which have better cellular coverage than rural areas. Finally, it was hypothesized that performance might be best during hours outside of peak cellular periods (i.e. outside of heavy commute times) since the cellular network might be overburdened during the commute hours. There were no hypotheses regarding performance by button push, weather, and in-vehicle unit — these analyses were exploratory.

2.2.5 RESULTS OF MOTOROLA'S USER GROUP DEPLOYMENT

In this section we present the results of each data analysis from the Motorola trials. In the following section (2.2.6) we present the results from the XYPOINT trials, followed by a section providing general conclusions drawn from the overall user group deployment test (2.2.7).

2.2.5.1 Motorola's Overall Success of Trials

Figure 2.2 presents the results of trials using the Motorola device over the life of the user group deployment. Only trials that resulted in a rating of Successful, Probably Successful, Probably Unsuccessful, or Unsuccessful are reported. (293 trials that resulted in conflicting data were discarded.) Of the 1,585 tests with one of these four ratings, about 88.5% of trials were Successful or Probably Successful, while about 11.5% of trials were Unsuccessful or Probably Unsuccessful.

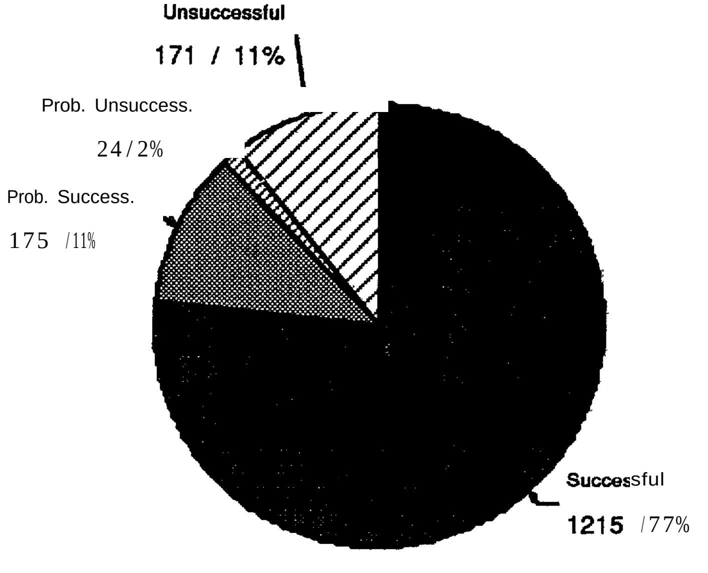


Figure 2.2 - Motorola User Group Deployment: Successful and Unsuccessful Trials

2.2.5.2 Motorola Time to Complete Trial

Two time measurements were taken for analysis of time to complete trial: (1) the user reported time from button push to time of call acknowledgement at the CSC, and (2) the computer-generated time from call receipt at the CSC to time call was verified.

2.2.5.2.1 User Reported Time from Button Push to Call Acknowledgment

For each trial, volunteers indicated on their user response form the time when they pushed the button and, for successful or probably successful calls, the time when the call was verbally acknowledged by the CSC operator. Volunteers usually reported time to the minute (rather than second). The results are presented in Figure 2.3.

Of the 1,390 trials that resulted in a rating of Successful or Probably Successful 1,250 (90%) produced user forms with usable time data. Of these, 1,151 (92%) were said to be acknowledged within three minutes and 418 (33%) within one minute.

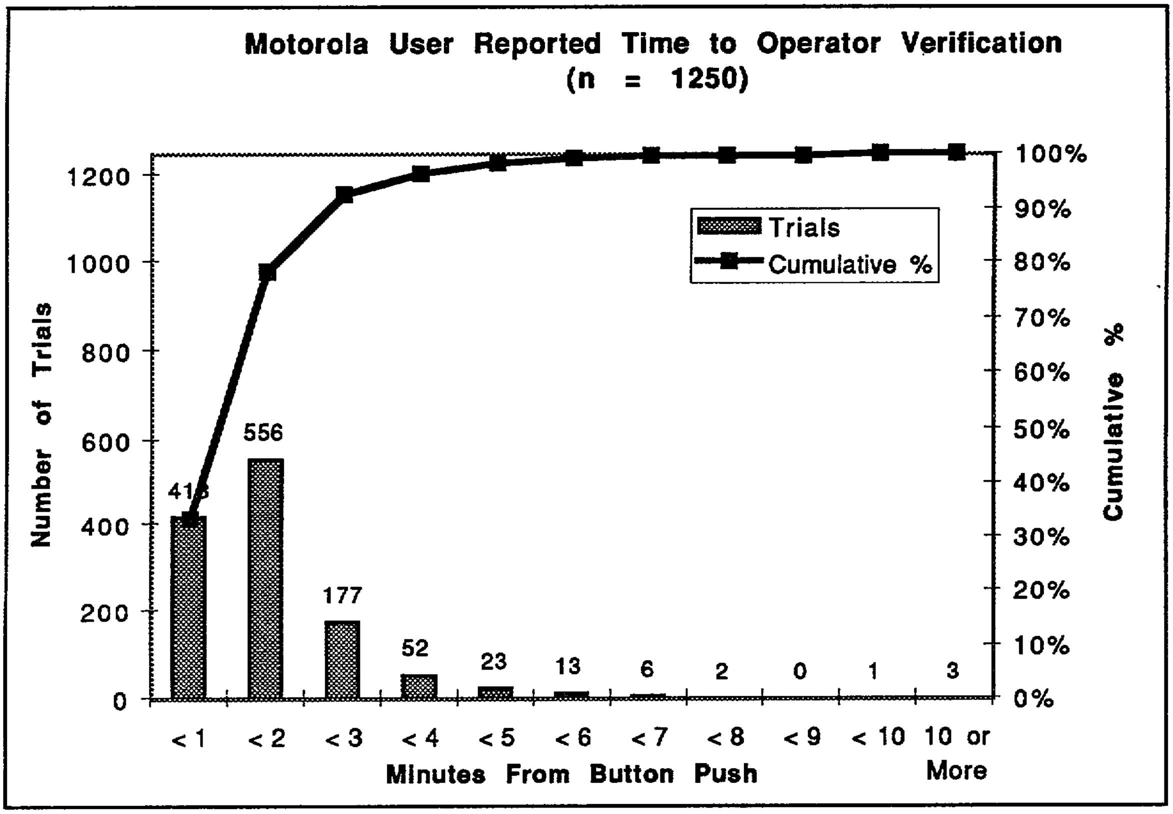


Figure 2.3 - Motorola User-Reported Time to Operator Acknowledgement

2.2.5.2.2 Computer Generated Data - Time from Call Receipt at CSC to Call Verification

For each trial, the call was timestamped when first received at the CSC . In addition, once an operator spoke with the caller, he or she indicated in the data that the call was verified. A timestamp was associated with this verification. Figure 2.4 presents the results of the time from call receipt at the CSC to call verification.

Of the 1,390 trials that resulted in a rating of Successful or Probably Successful, 1,044 (75%) produced usable time data in the CSC data stream. Of these, 1,010 (97%) were verified within three minutes and 764 (73%) were verified within one minute.

Since there was CSC data for each trial classified as Successful or Probably Successful (see Table 2.1), it seems surprising that only 75% of these trials provided usable CSC time data. (Compare with 97% producing usable time data for the comparable XYPOINT trials.) The vast majority of the data streams with unusable time data (288) were associated with multiple call IDs, indicating either that the call was dropped and incorrectly treated as a new call when reconnected, or that more than one call was incorrectly placed during the day from an individual unit. For these calls, the time differences were not included in this analysis since each call would have had to be subjectively evaluated to determine if the call was an unsuccesful reconnection of the same trial or a separate call altogether. (For further information on the dropped call issue, see discussion of the Dropped Carrier Test, Section 2.4.)

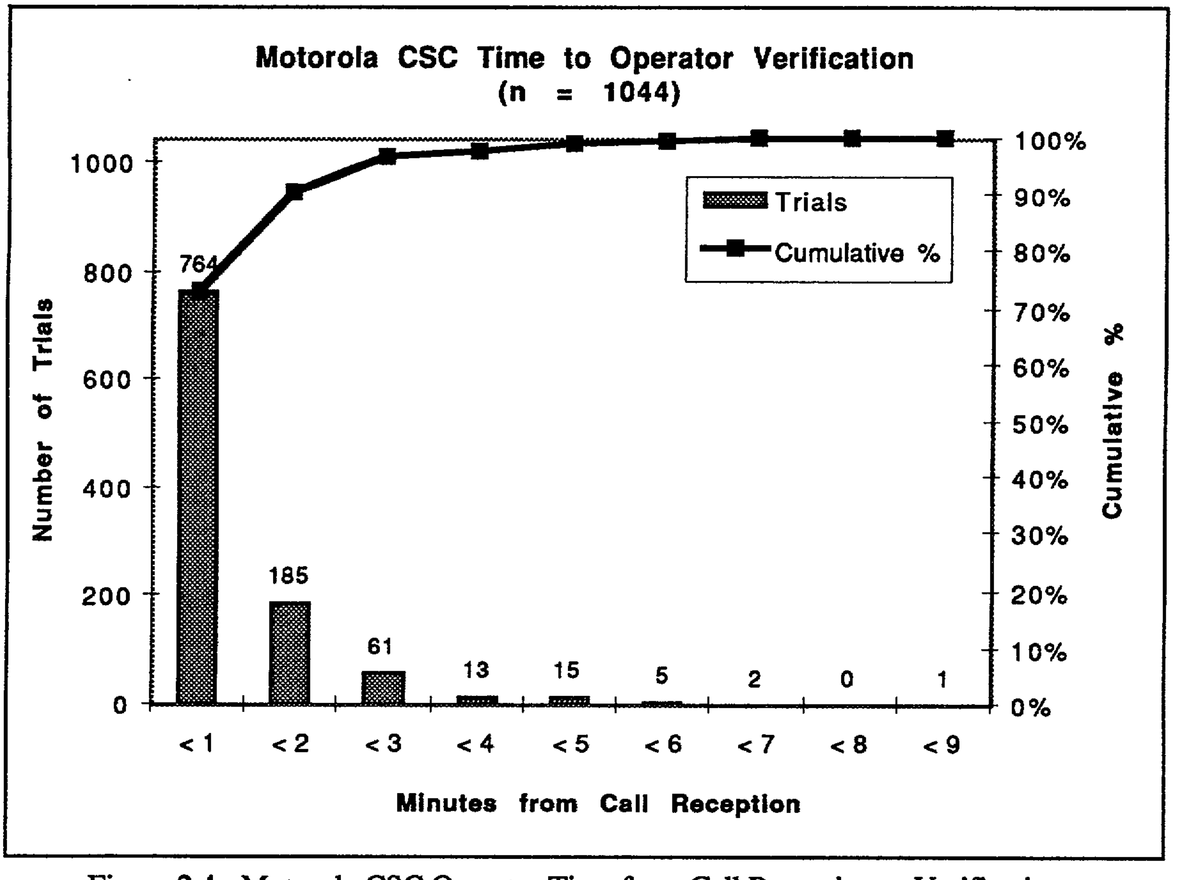


Figure 2.4 - Motorola CSC Operator Time from Call Reception to Verification

2.2.5.3 Motorola - Relationship Between Trial Success and Month of Testing

Figures 2.5 and 2.6 present trial ratings by month of the user group deployment. Figure 2.5 presents the *number* of successful and unsuccessful trials by month, while Figure 2.6 presents the *percent* of successful and unsuccessful trials by month. A Kruskal-Wallis one-way analysis of variance of ratings by month revealed that there was a statistically significant difference in ratings by month ($H = 39.20$; df = 5; $p < .0001$). Performance steadily improved from November (74% successful or probably successful) until February (94% successful or probably successful), with February being the most successful month of testing. March and April did not show this improvement.

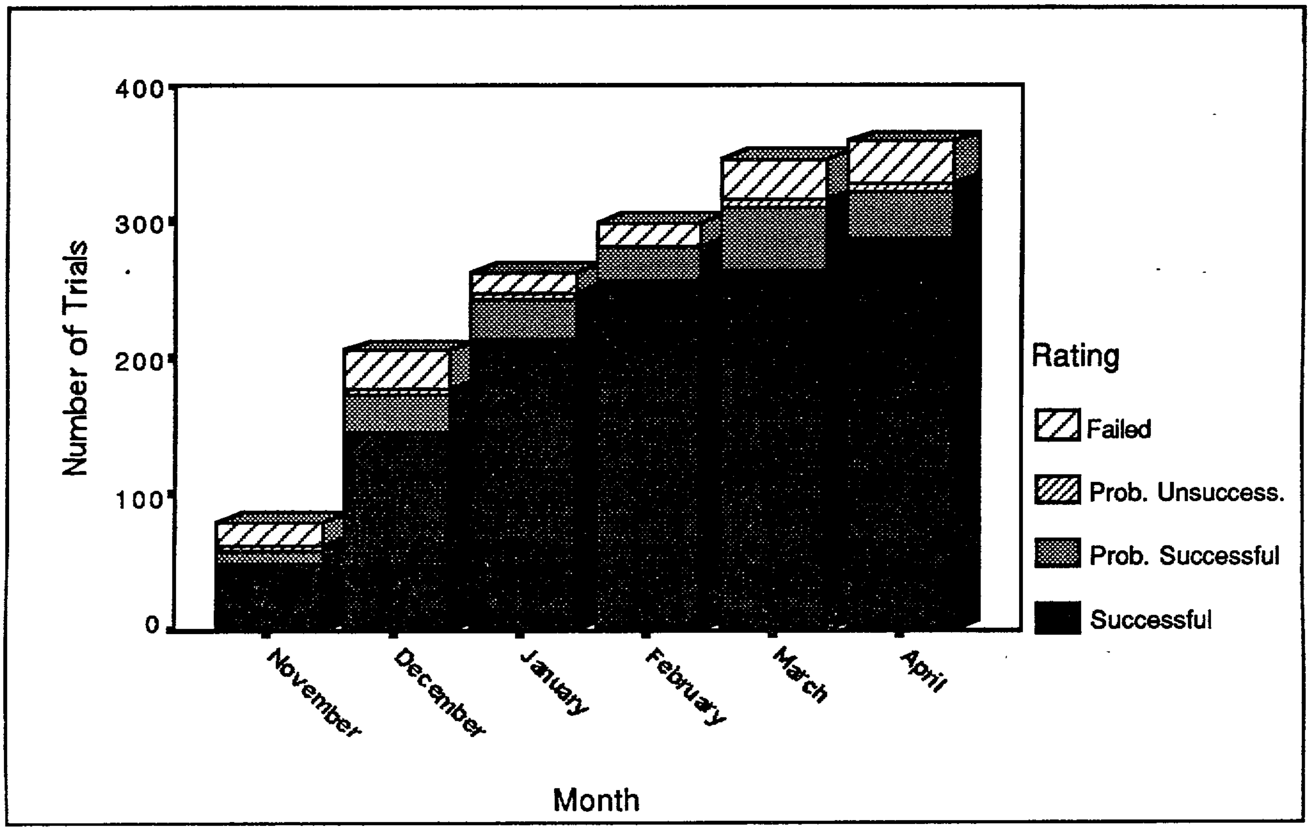


Figure 2.5 - Motorola Trial Success by Month (Number)

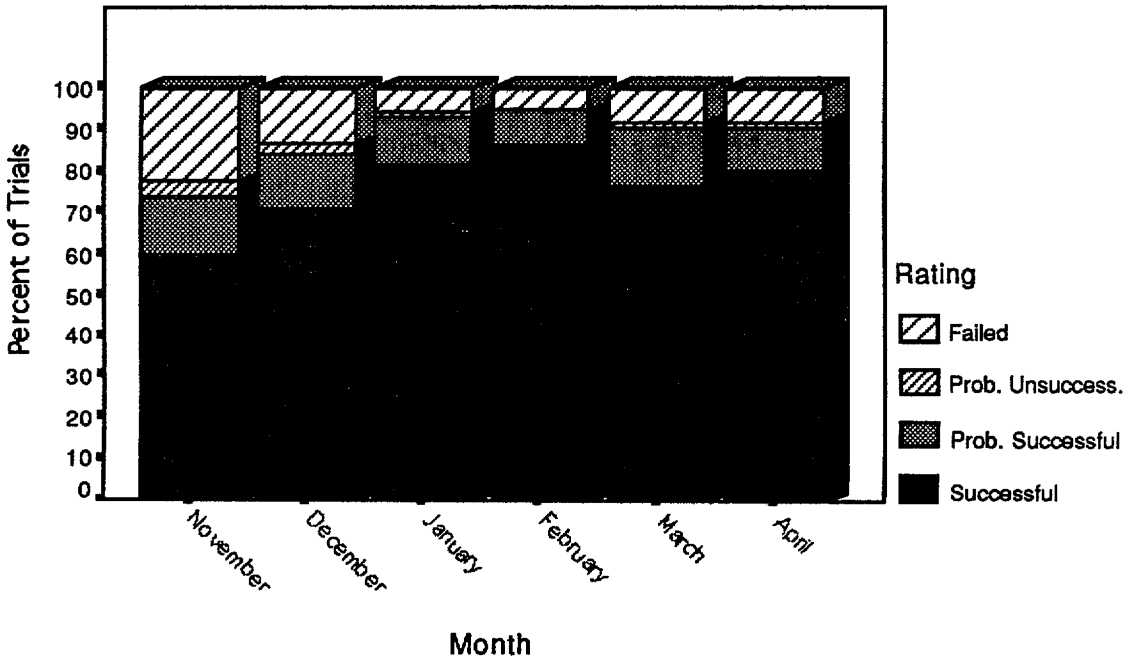


Figure 2.6 - Motorola Trial Success by Month (%)

2.2.5.4 Motorola - Relationship Between Trial Success and Time of Day

Figures 2.7 and 2.8 present trial ratings by time of day between 2 P.M. and 6 P.M. Figure 2.7 presents the *number* of successful and unsuccessful trials by time of day, while Figure 2.8 presents the *percent* of successful and unsuccessful trials by time of day. A Kruskal-Wallis one-way analysis of variance of ratings by time of day revealed that there was a statistically significant difference in ratings by time of day ($H = 24.02$; $df = 7$; $p < .01$). Testing was most successful between 4:30 - 4:59 (92% successful or probably successful) and least successful between 2 - 2:29 (75% successful or probably successful).

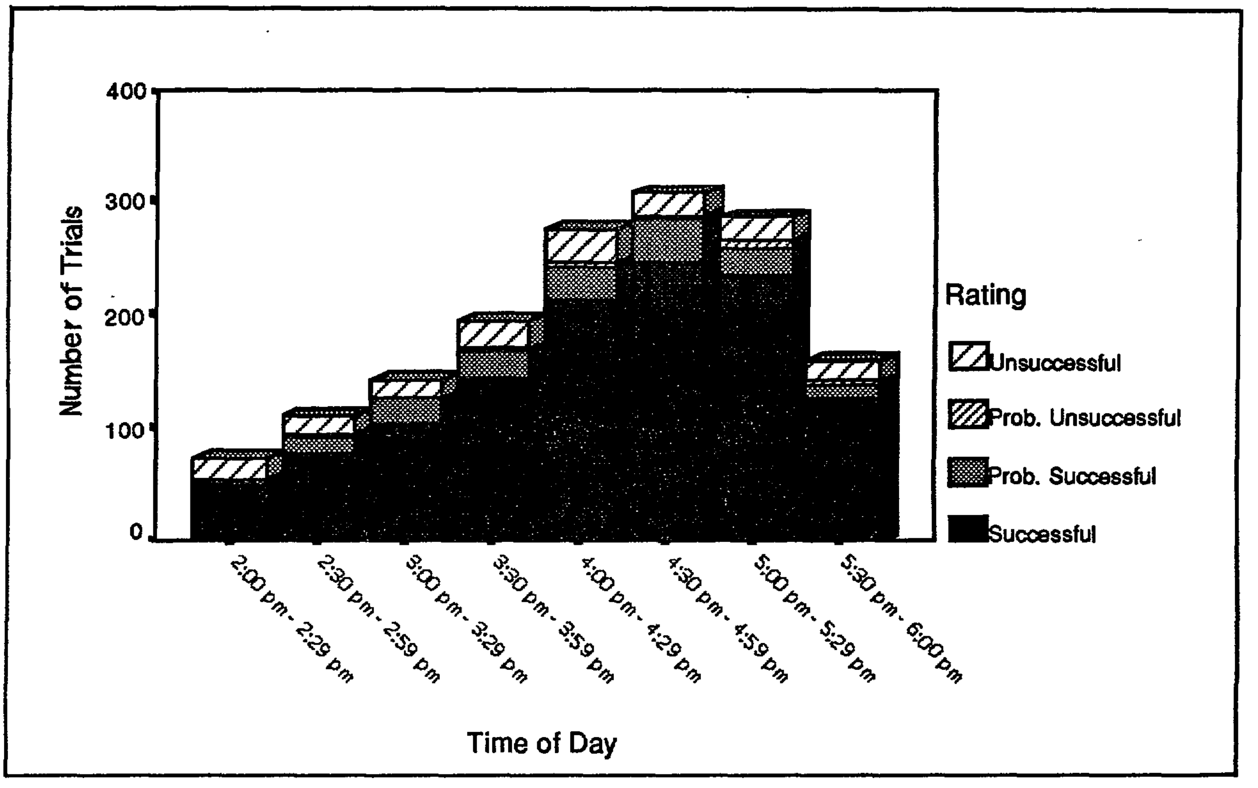


Figure 2.7 - Motorola Trial Success by Time of Day (Number)

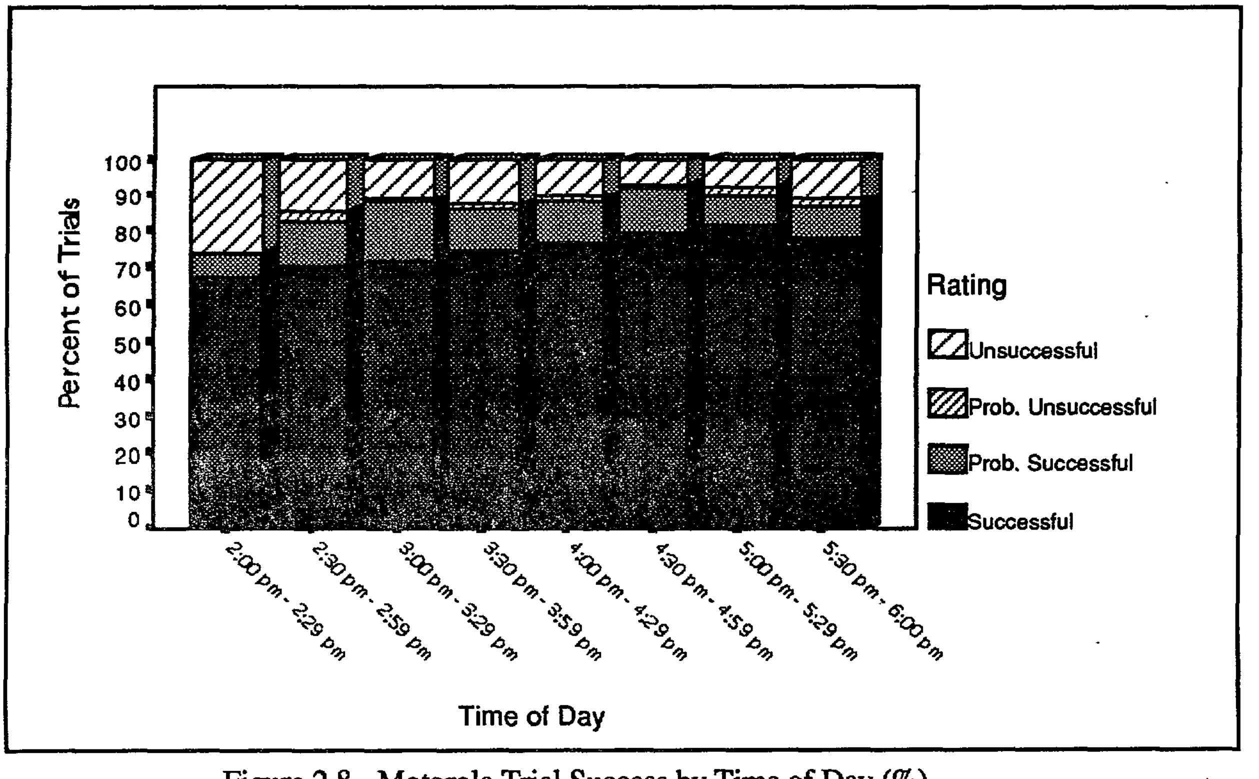


Figure 2.8 - Motorola Trial Success by Time of Day (%)

2.2.5.5 Motorola - Relationship Between Trial Success and Button Pushed

Figure 2.9 presents trial ratings by button pushed. A Kruskal-Wallis one-way analysis of variance of ratings by button pushed revealed no statistically significant difference in ratings. Statistically, the system performed equally well for each button pushed (emergency - 88% successful or probably successful; roadside assistance - 90% successful or probably successful; traveler assistance - 87% successful or probably successful).

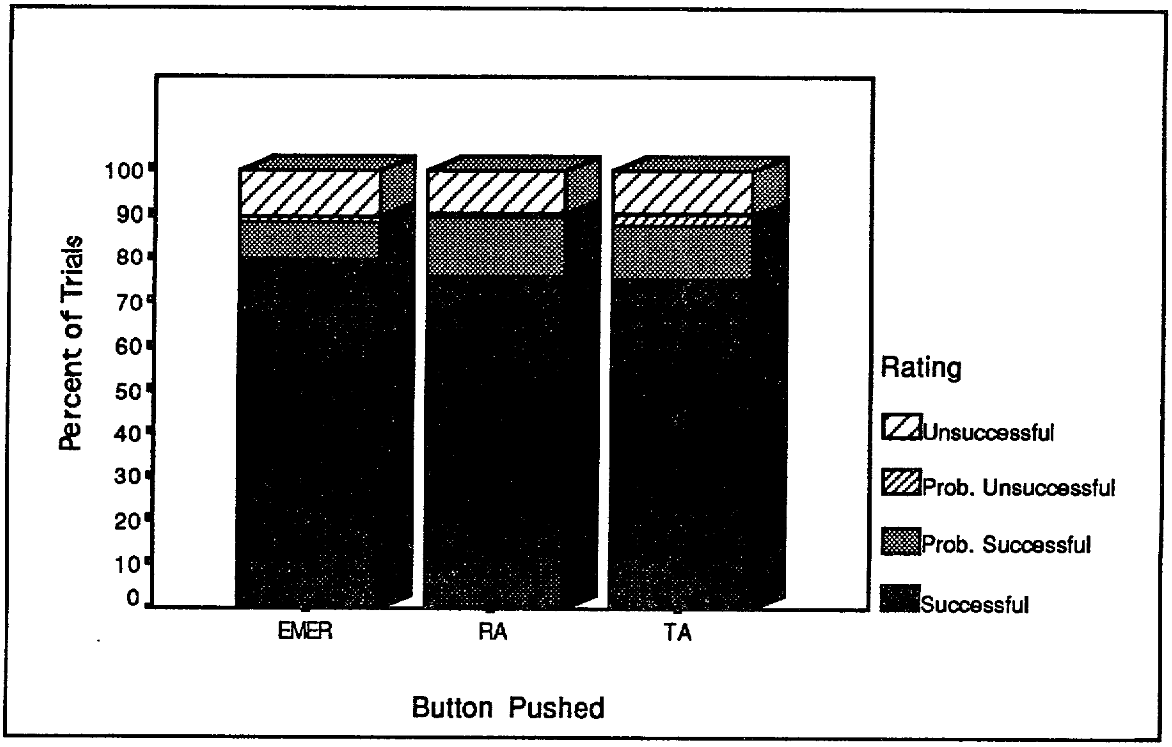


Figure 2.9 - Motorola Trial Success by Button Pushed

2.2.5.6 Motorola - Relationship Between Trial Success and User Reported Weather Conditions

Figures 2.10 and 2.11 present trial ratings by user reported weather conditions. Figure 2.10 presents the *number* of successful and unsuccessful trials by weather, while Figure 2.11 presents the *percent* of successful and unsuccessful trials by weather. A Kruskal-Wallis one-way analysis of variance of ratings by weather revealed no statistically significant difference in ratings by weather. Statistically, the system performed equally well in cloudy, rainy, snowy, or clear conditions (overcast - 89% successful or probably successful; rain - 89% successful or probably successful; snow - 91% successful or probably successful; clear - 88% successful or probably successful).

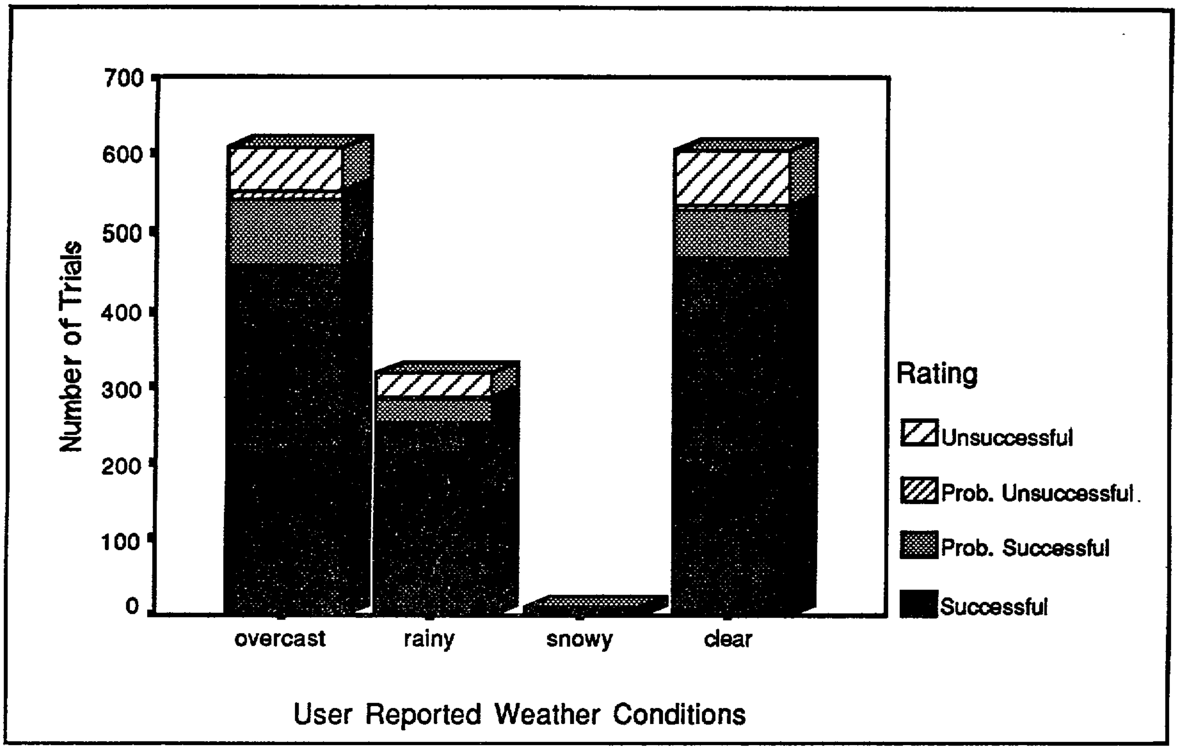


Figure 2.10 - Motorola Trial Success by Weather (Number)

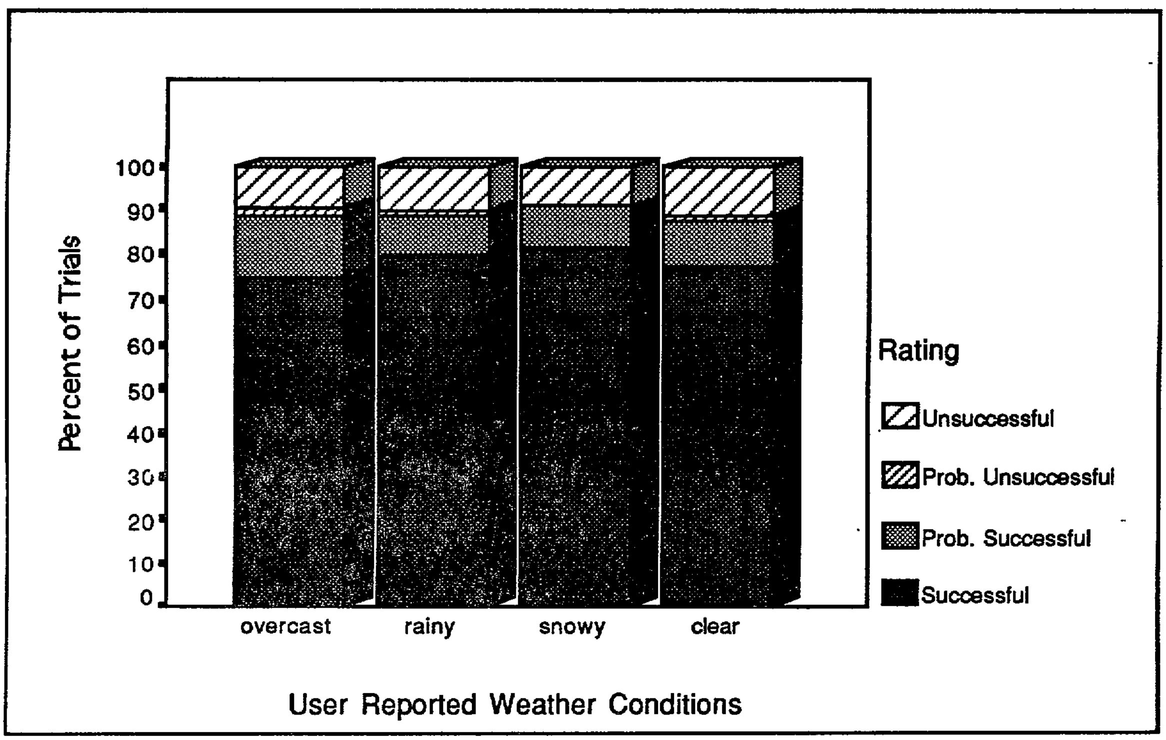


Figure 2.11 - Motorola Trial Success by Weather (%)

2.2.5.7 Motorola - Relationship Between Trial Success and User Reported Trial Location

Figures 2.12 and 2.13 present trial ratings by user reported locations. Figure 2.12 presents the *number* of successful and unsuccessful trials by location, while Figure 2.13 presents the *percent* of successful and unsuccessful trials by location. A Kruskal-Wallis one-way analysis of variance of ratings by location revealed no statistically significant difference in ratings by location. Statistically, the system performed equally well in urban, suburban, or rural areas (urban - 89% successful or probably successful; suburban - 90% successful or probably successful; rural - 88% successful or probably successful).

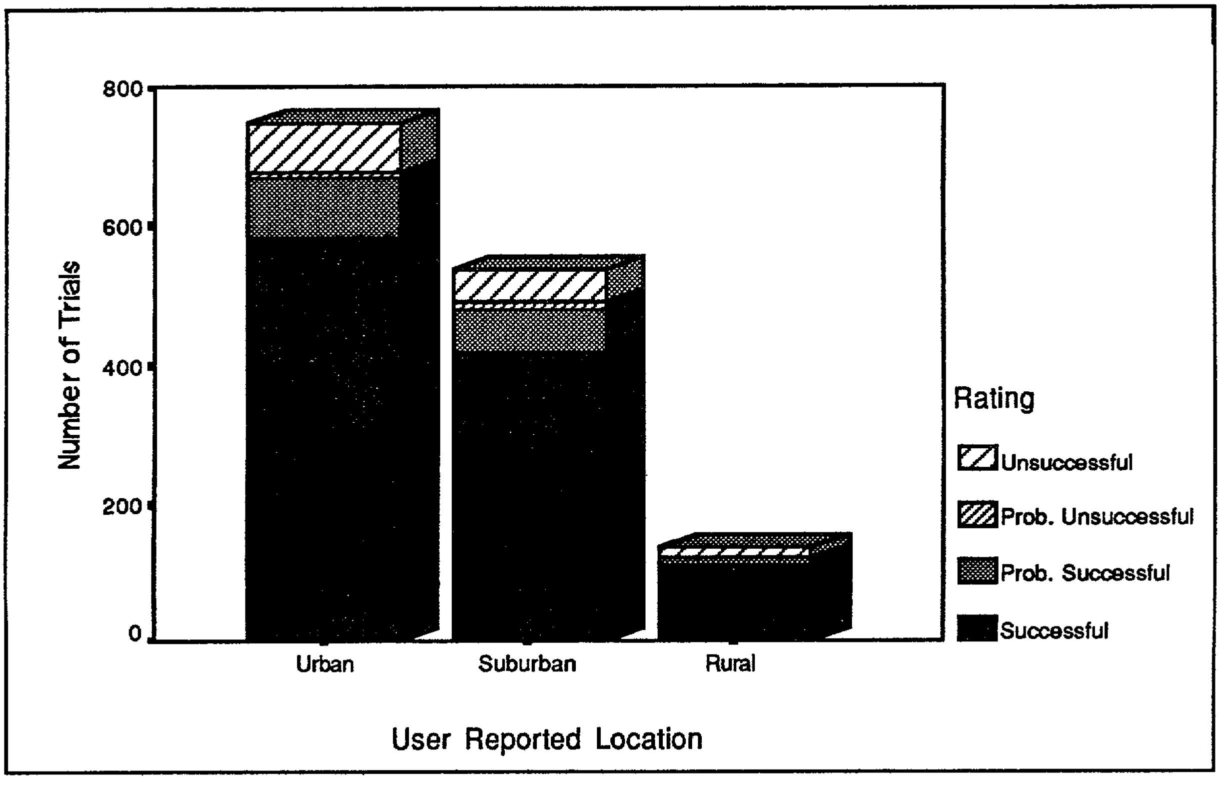


Figure 2.12 - Motorola Trial Success by Location (Number)

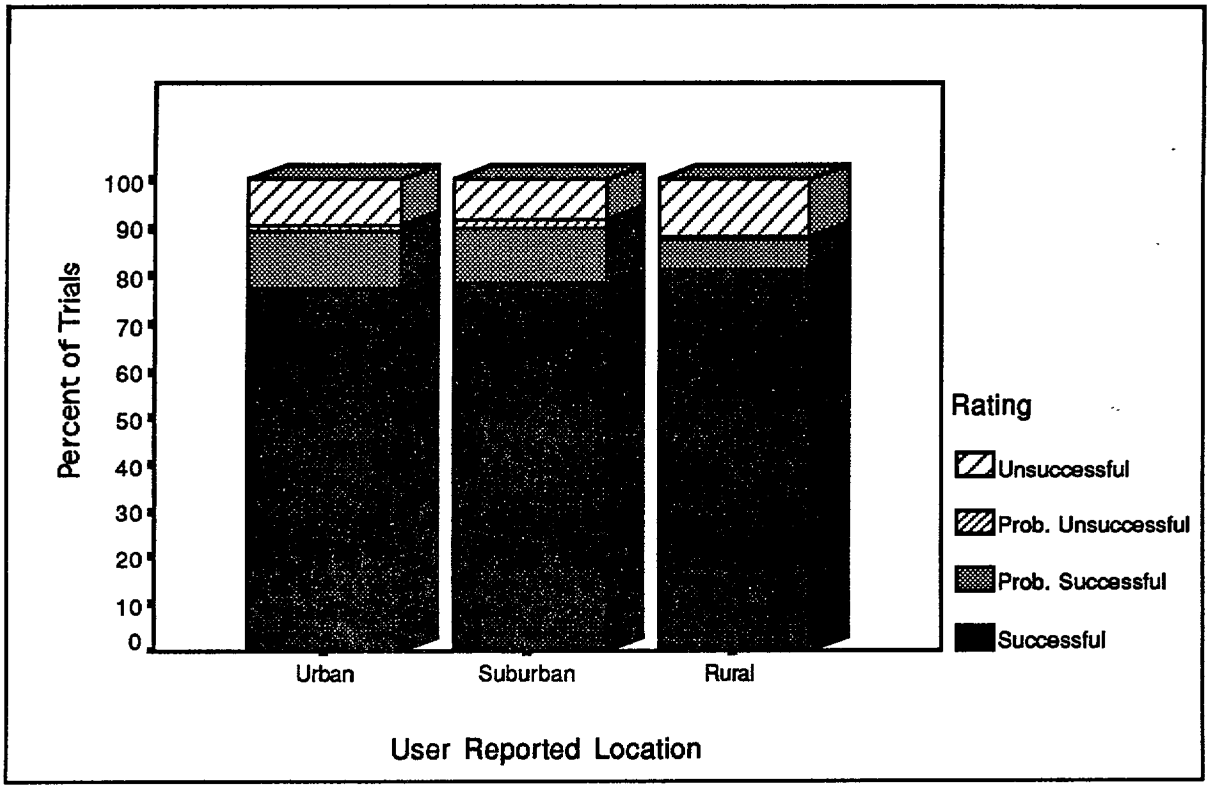


Figure 2.13 - Motorola Trial Success by Location (%)

2.2.5.8 Motorola - Relationship Between Trial Success and Individual In-Vehicle Unit

Figures 2.14 and 2.15 present the count and percentage of trial ratings by in-vehicle unit. Figure 2.14 presents the *number* of successful and unsuccessful trials by unit, while Figure 2.15 presents the *percent* of successful and unsuccessful trials by unit. A Kruskal-Wallis one-way analysis of variance of ratings by unit revealed a statistically significant difference (H= 101.93; df = 38; p < .0001). Some units performed better than others. In general, the units that did not function well (e.g.units 10, 13, and 18) took part in fewer trials, perhaps indicating that they were replaced.

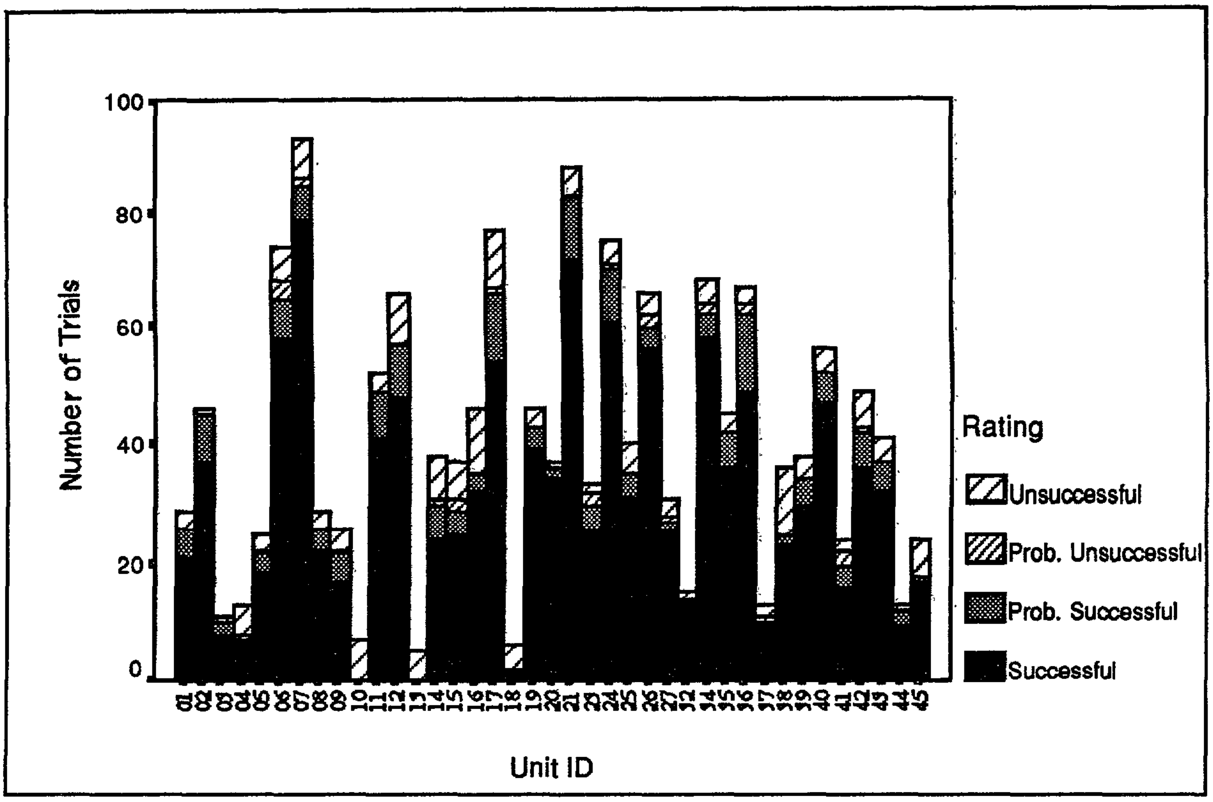


Figure 2.14 - Motorola Trial Success by Unit (Number)

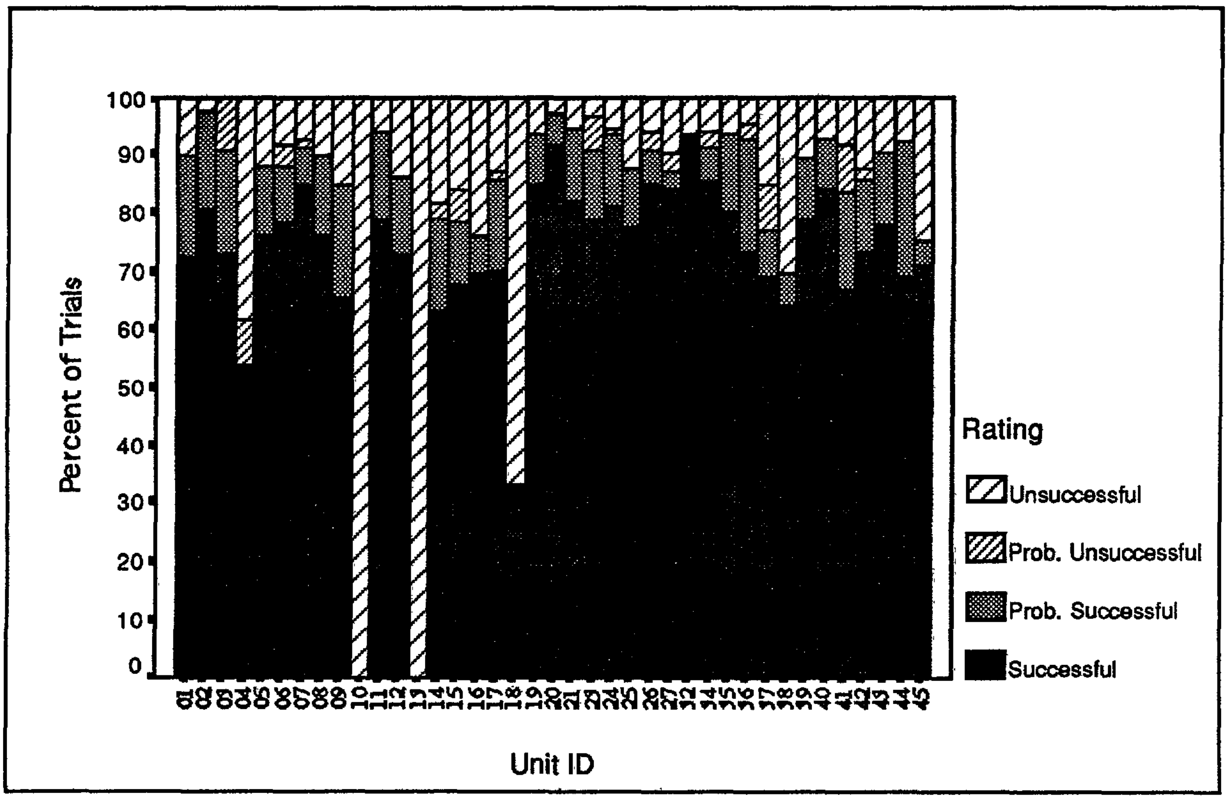


Figure 2.15 - Motorola Trial Success by Unit (%)

2.2.6 RESULTS OF XYPOINT'S USER GROUP DEPLOYMENT

In this section we present the results of each data analysis from the XYPOINT trials. This section is followed by a section providing general conclusions drawn from the overall user group deployment test (2.2.7).

2.2.6.1 XYPOINT's Overall Success of Trials

Figure 2.16 presents the results of the 3,123 XYPOINT trials that resulted in a rating of Successful, Probably Successful, Probably Unsuccessful, or Unsuccessful. Over the life of the user group deployment, about 66% of trials were successful or probably successful, while about 34% of trials were unsuccessful or probably unsuccessful.

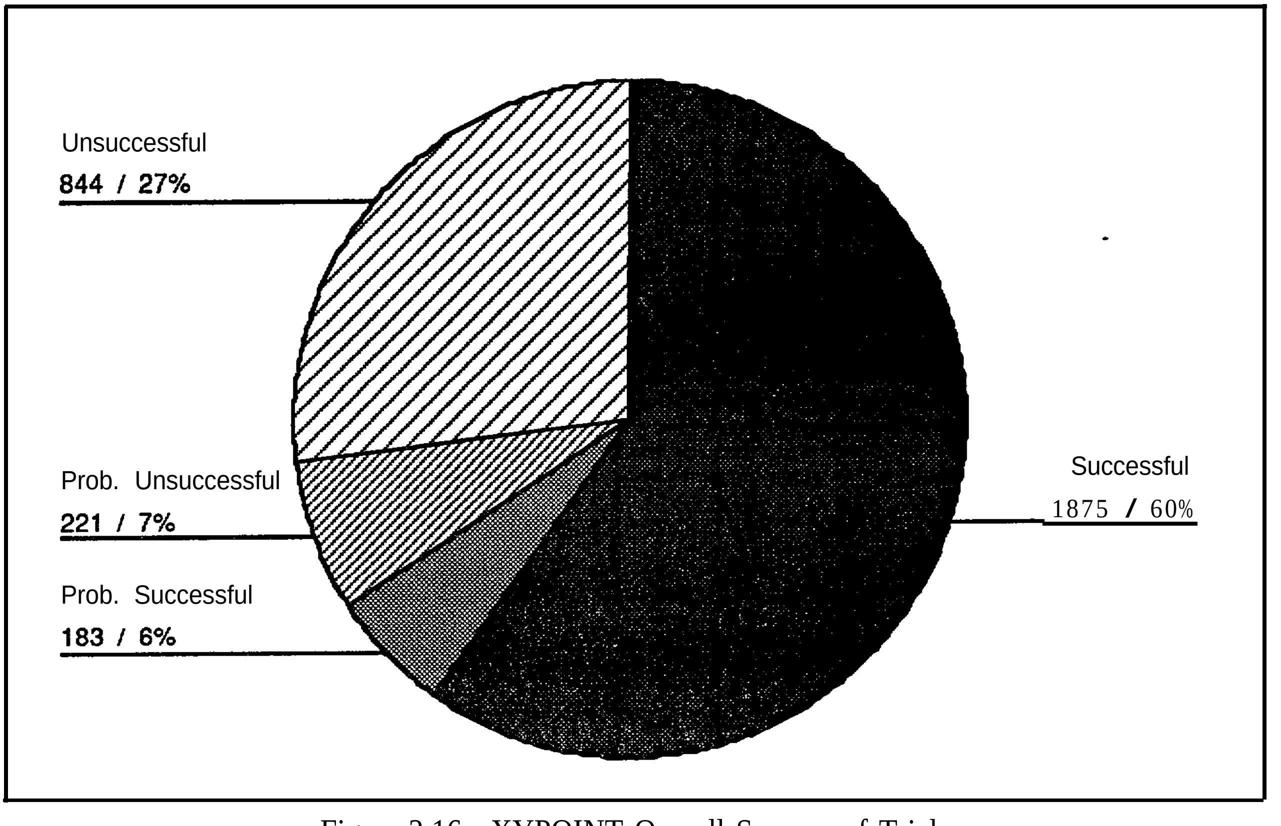


Figure 2.16 - XYPOINT Overall Success of Trials

2.2.6.2 XYPOINT User Group Deployment Results - Time to Complete Trial

Two time measurements were taken for this analysis: (1) the user reported time from button push to the time a message was displayed on the in-vehicle unit indicating the trial was completed, and (2) the computer-generated time from call receipt at the CSC to the time a "Yes" or "No" response was received from the user.

2.2.6.2.1 User Reported Time from Button Push to Trial Completion

For each trial, volunteers were to indicate on their user response form the time when they pushed the button and the time when "PUSHME" was displayed on the in-vehicle device, indicating a successful completion of the trial. Volunteers usually reported time to the minute (rather than second). The results are presented in Figure 2.17.

Of the 2,058 trials that resulted in a rating of Successful or Probably Successful, 1,982 (96%) produced user forms with usable time data. Of these, 1,634 (82%) were said to be acknowledged within three minutes and 508 (26%) within one minute.

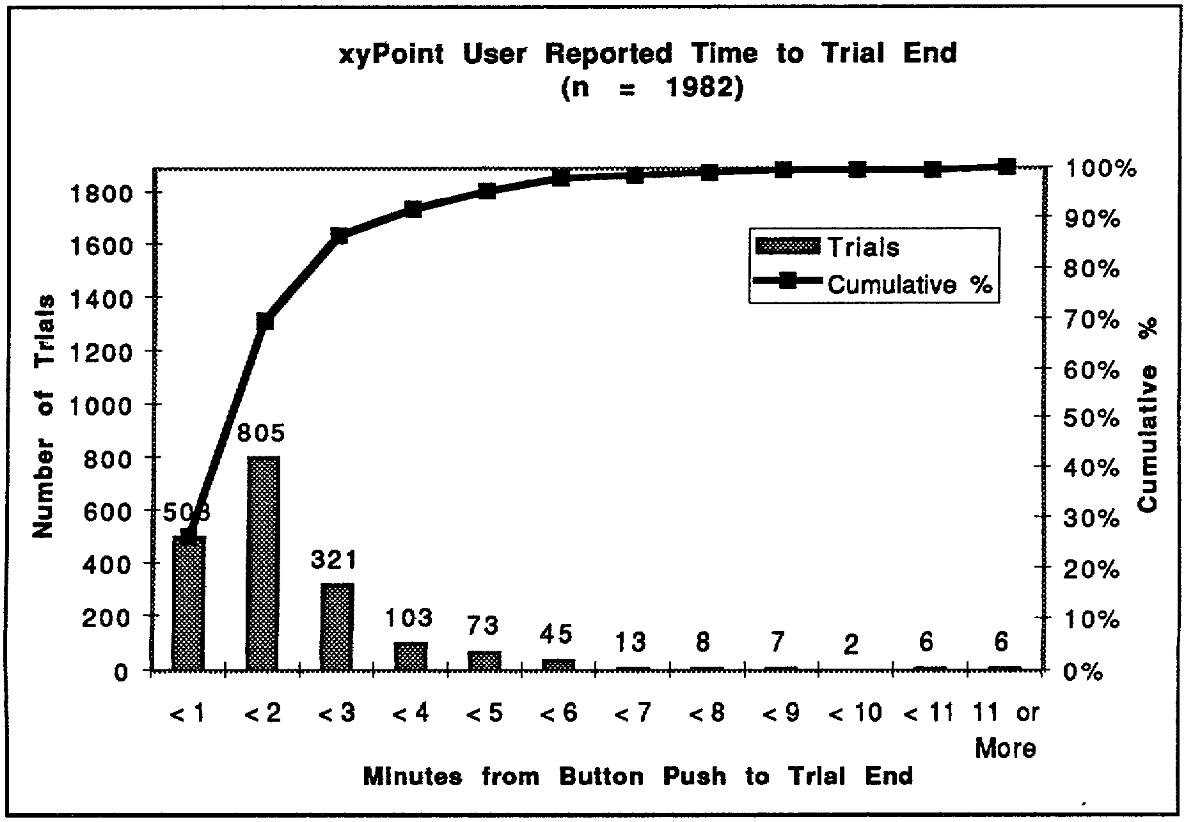


Figure 2.17 XYPOINT User-Reported Time from Button Push to Trial End

2.2.6.2.2 Computer Generated Data - Time from Call Receipt at CSC to Call Verification

For each trial, the reception of a call at the CSC was automatically timestamped. After the call was received, the operator sent a message back asking the user to confirm the type of call. The user confirmed by pressing the "Yes" or "No" button on the in-vehicle device. When the "Yes" or "No" was received at the CSC, this reception was also automatically timestamped. For the XYPOINT system, we considered the reception of "Yes" or "No" to be verification of the call. Figure 2.18 presents the results of the time from call receipt at the CSC to call verification.

Of the 2,058 trials that resulted in a rating of Successful or Probably Successful, 1,994 (97%) produced usable time data in the CSC data stream. Of these, 1,965 (98.5%) were verified within three minutes and 1,858 (93%) were verified within one minute.

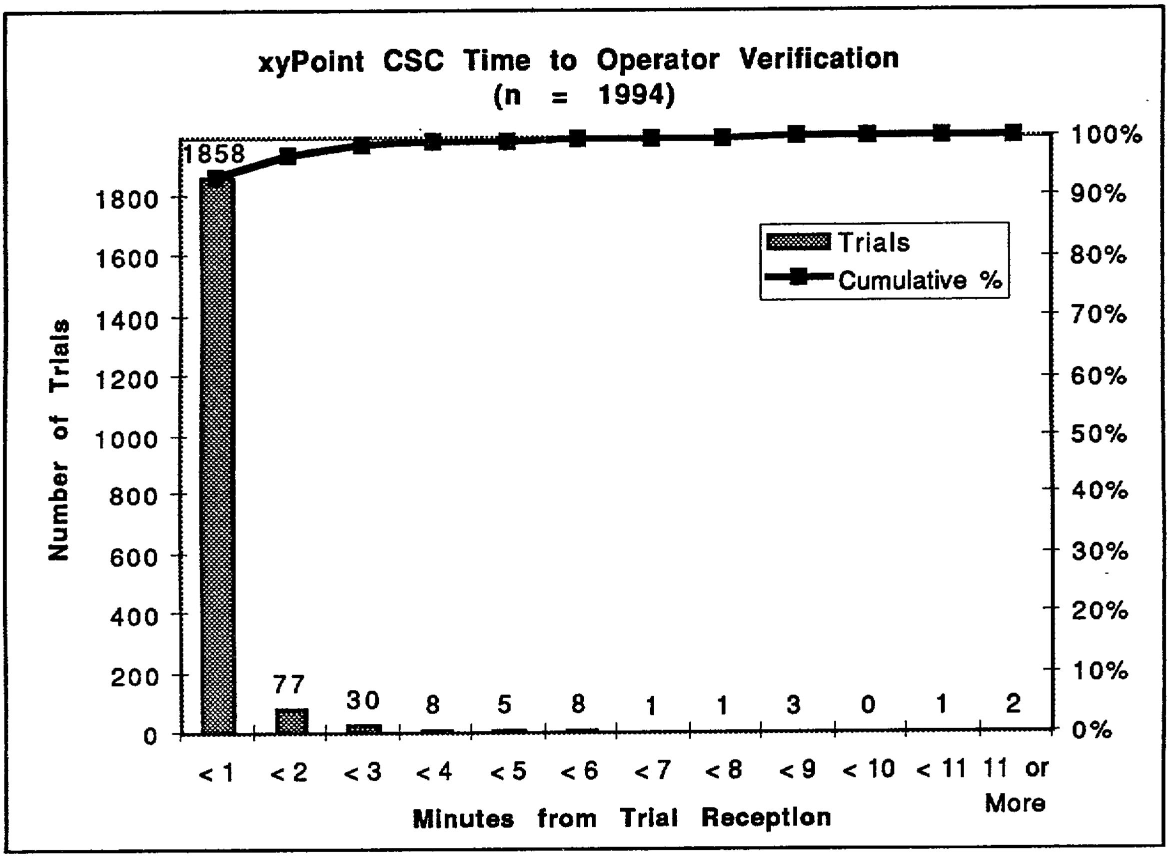


Figure 2.18 - XYPOINT CSC Operator Time from Call Reception to Verification

2.2.6.3 XYPOINT - Relationship Between Trial Success and Month of Testing

Figures 2.19 and 2.20 present trial ratings by month of the user group deployment. Figure 2.19 presents the *number* of successful and unsuccessful trials by month, while Figure 2.20 presents the *percent* of successful and unsuccessful trials by month. A Kruskal-Wallis one-way analysis of variance of ratings by month revealed that there was a statistically significant difference in ratings by month ($H = 93.14$; df = 4; $p < .01$). There was a trend towards improved success as the test went on with 51% successful or probably successful at the beginning of the test and 74% successful or probably successful towards the end. This trend was most likely due to the continued improvement in the wireless CDPD environment which was still in its infancy during the test.

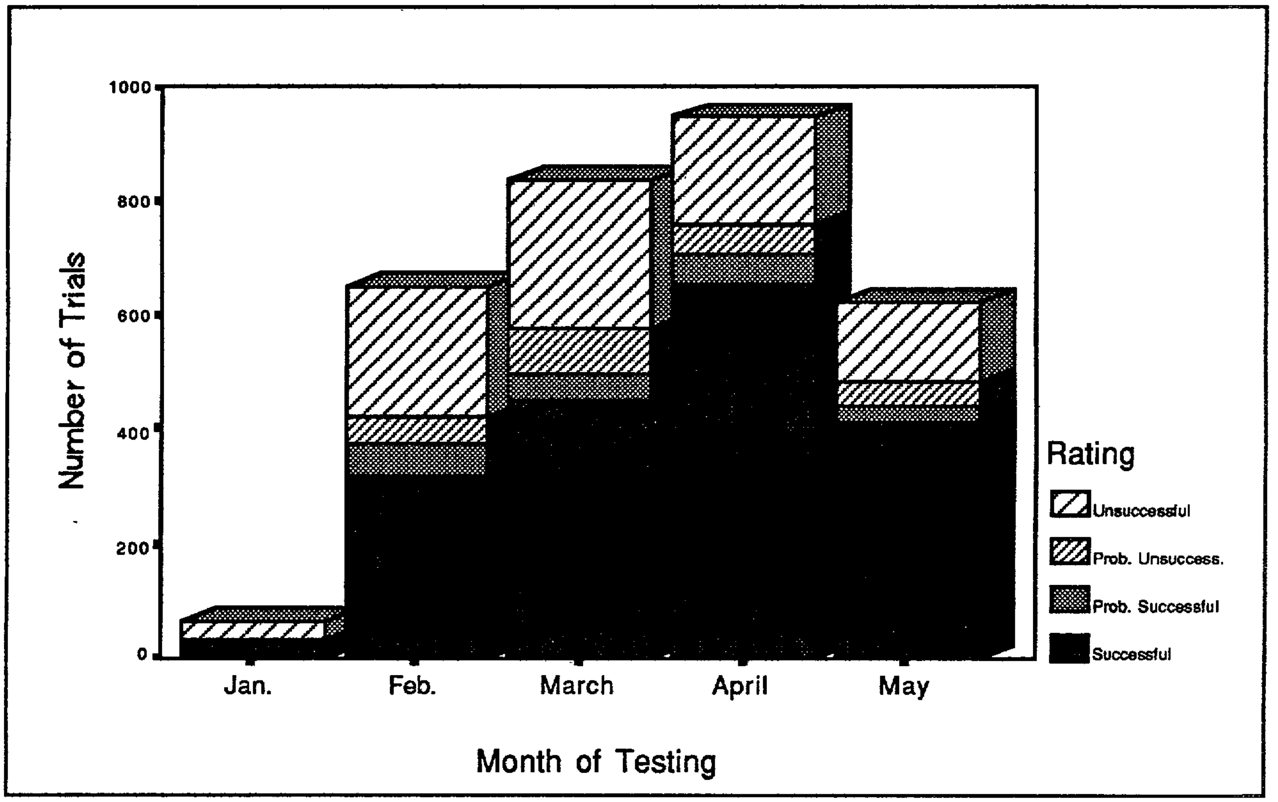


Figure 2.19 - XYPOINT Trial Success by Month (Number)

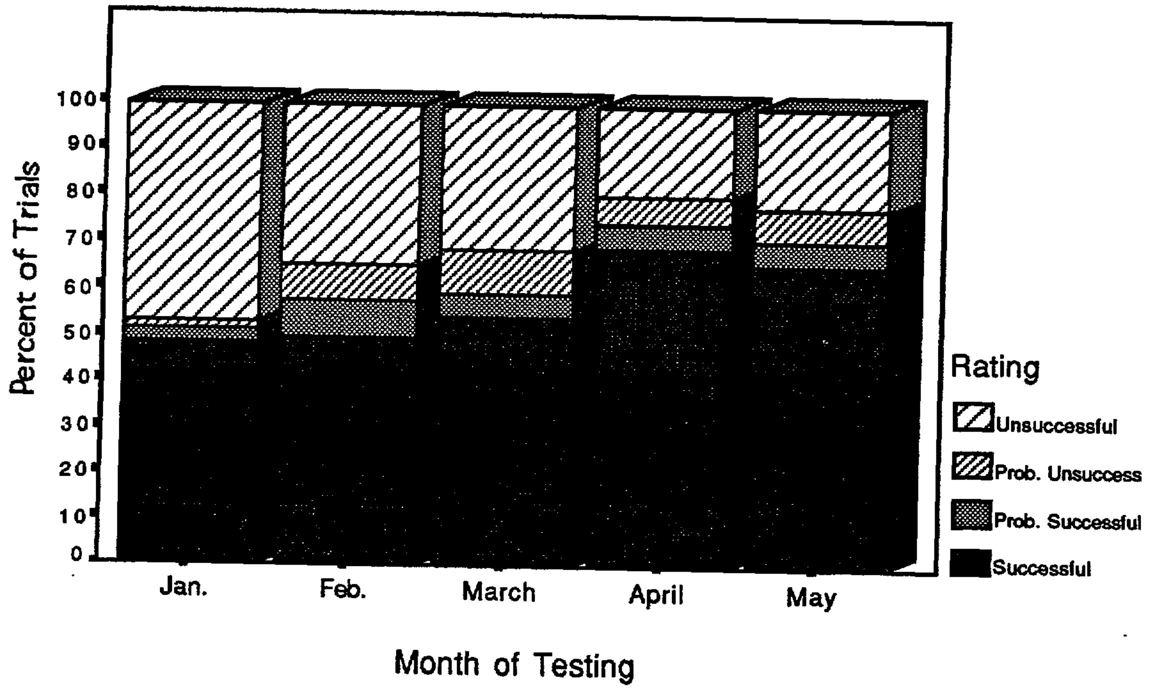


Figure 2.20 - XYPOINT Trial Success by Month (%)

2.2.6.4 XYPOINT - Relationship Between Trial Success and Time of Day

Figures 2.21 and 2.22 present trial ratings by time of day between 7 AM and 6 PM. Figure 2.21 presents the *number* of successful and unsuccessful trials by time of day, while Figure 2.22 presents the *percent* of successful and unsuccessful trials by time of day. A Kruskal-Wallis one-way analysis of variance of ratings by time of day indicated a statistically significant difference in trial ratings by time of day ($H = 27.19$; df = 10; $p < .01$), indicating that the performance differed depending on the hour of the day. System performance was least successful between 5 and 5:59 PM (about 35% Unsuccessful or Probably Unsuccessful) and between 7 and 7:59 AM (about 32% Unsuccessful or Probably Unsuccessful). This was most likely due to CDPD being forced off the air during peak cellular usage. (This has since been corrected in the AT&T CDPD based systems by using CDPD dedicated channels unaffected by cellular usage.)

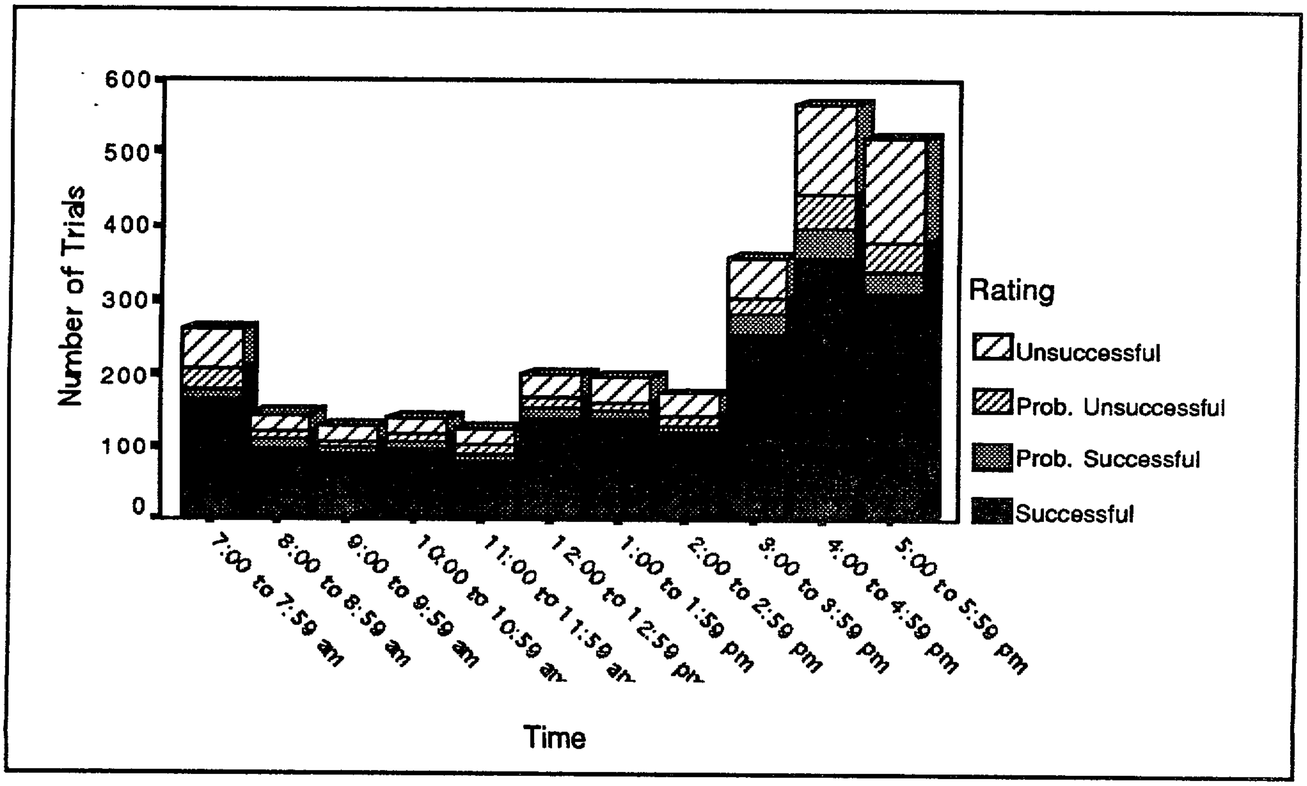


Figure 2.21 - XYPOINT Trial Success by Time of Day (Number)

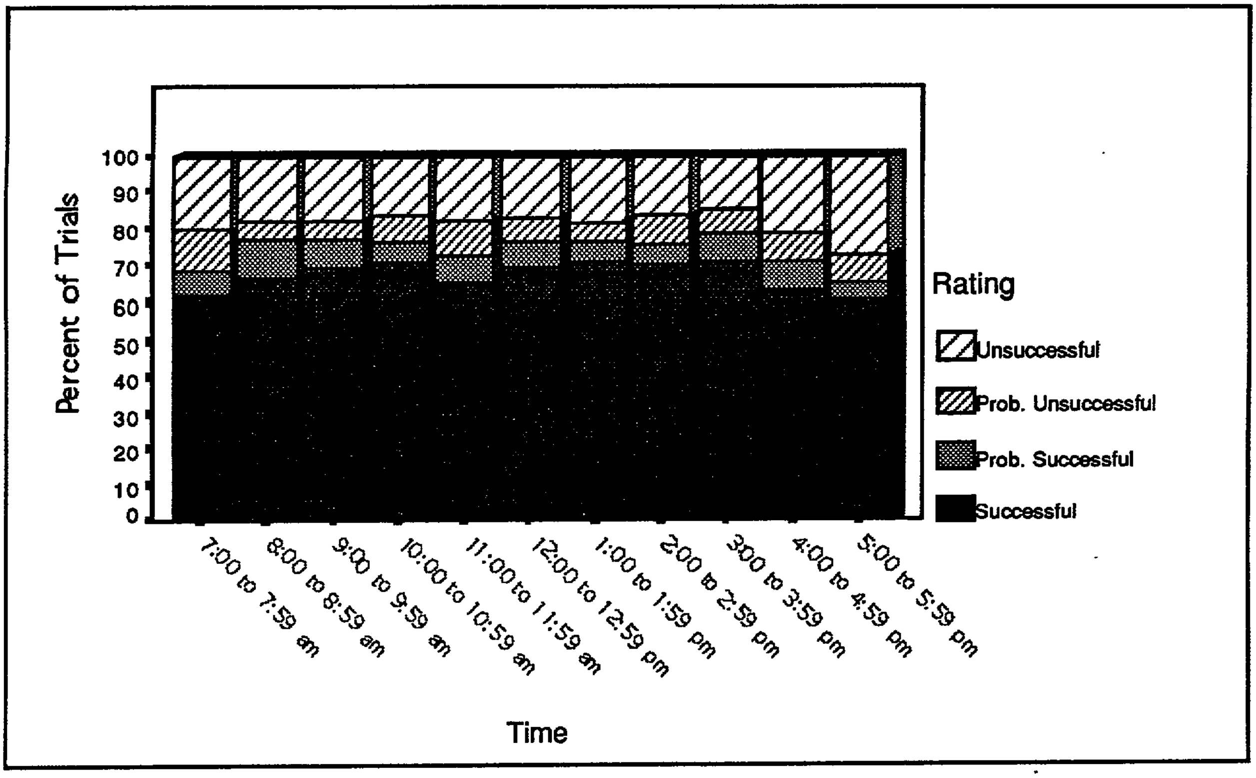


Figure 2.22 - XYPOINT Trial Success by Time of Day (%)

2.2.6.5 XYPOINT - Relationship Between Trial Success and Button Pushed

Figures 2.23 and 2.24 present trial ratings by button pushed. Figure 2.23 presents trial ratings by initial button pushed (911, Medical, or Auto), while Figure 2.24 presents trial results by the confirmation button pushed (Yes or No). A Kruskal-Wallis one-way analysis of variance of ratings by initial button pushed revealed no statistically significant difference in ratings. Similarly, a Kruskal-Wallis one-way analysis of variance of ratings by confirmation button pushed revealed no statistically significant difference in ratings.

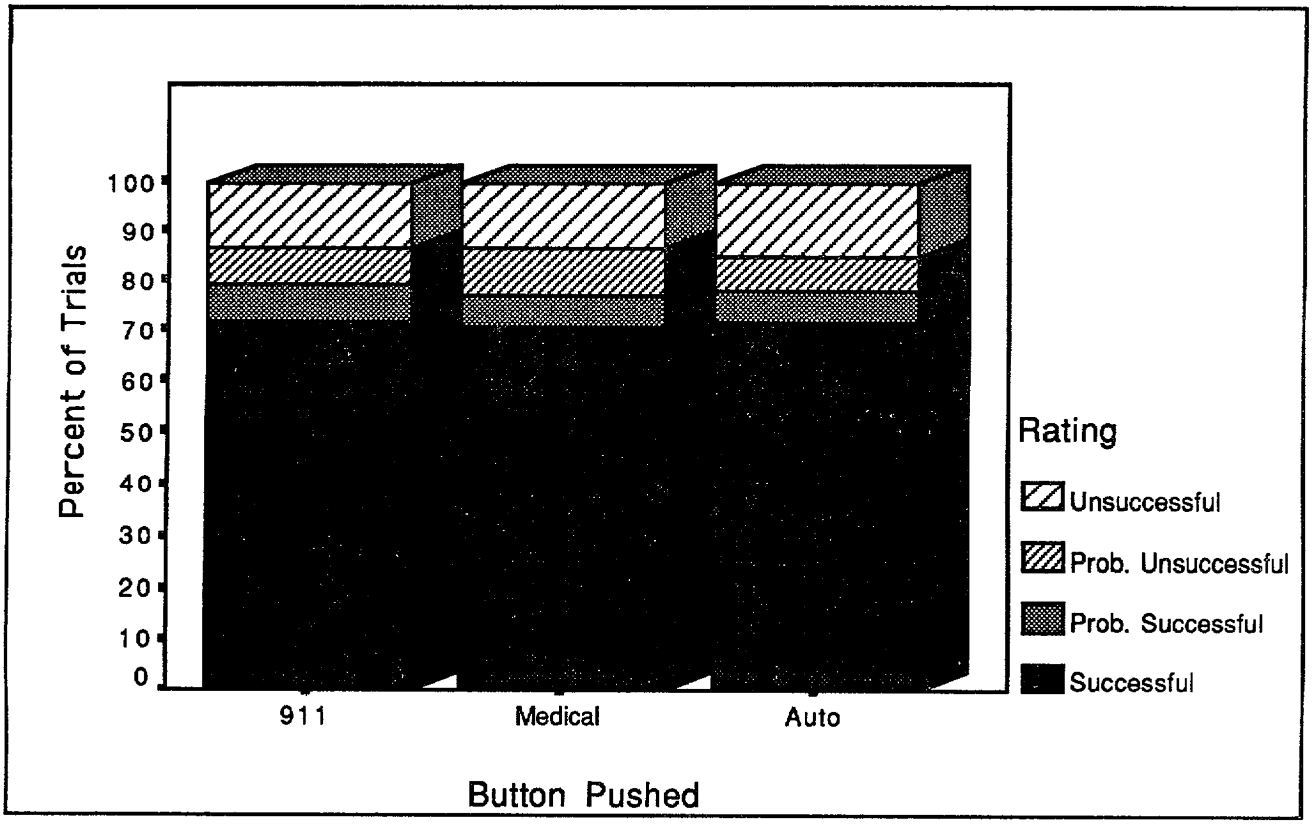


Figure 2.23 - XYPOINT Trial Success by Initial Button Pushed

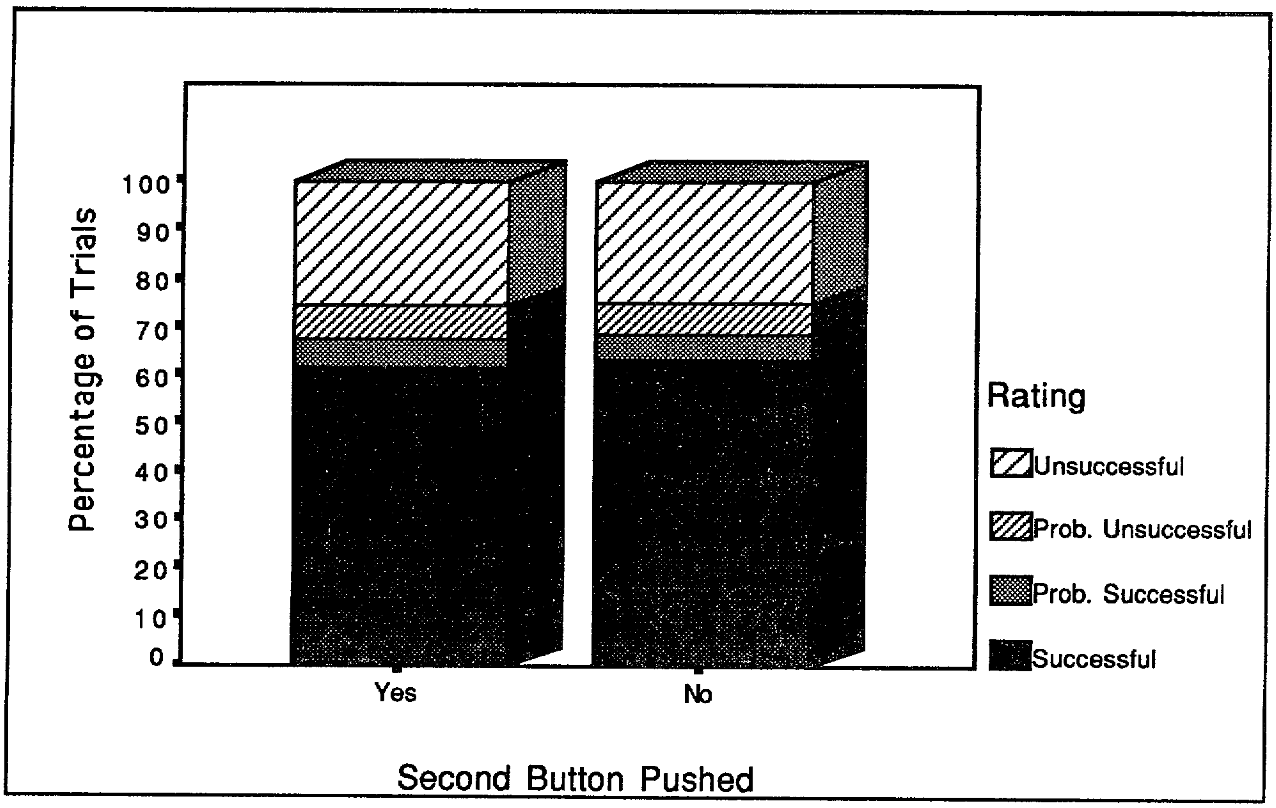


Figure 2.24 - XYPOINT Trial Success by Confirmation Button Pushed

2.2.6.6 XYPOINT - Relationship Between Trial Success and User Reported Weather Conditions

Figures 2.25 and 2.26 present percentage and number of successful trials by user reported weather conditions. Figure 2.25 presents the *number* of successful and unsuccessful trials by weather, while Figure 2.26 presents the *percent* of successful and unsuccessful trials by weather. A Kruskal-Wallis one-way analysis of variance of ratings by weather revealed a statistically significant difference in ratings by weather condition (H = 8.93; df = 3; p < .05). While these differences were statistically significant, they do not appear to be particularly meaningful (overcast - 69% successful or probably successful; rain - 73% successful or probably successful; snow - 74% successful or probably successful; clear - 68% successful or probably successful). In addition, it should be noted that there were an extremely small number of trials conducted under snowy conditions.

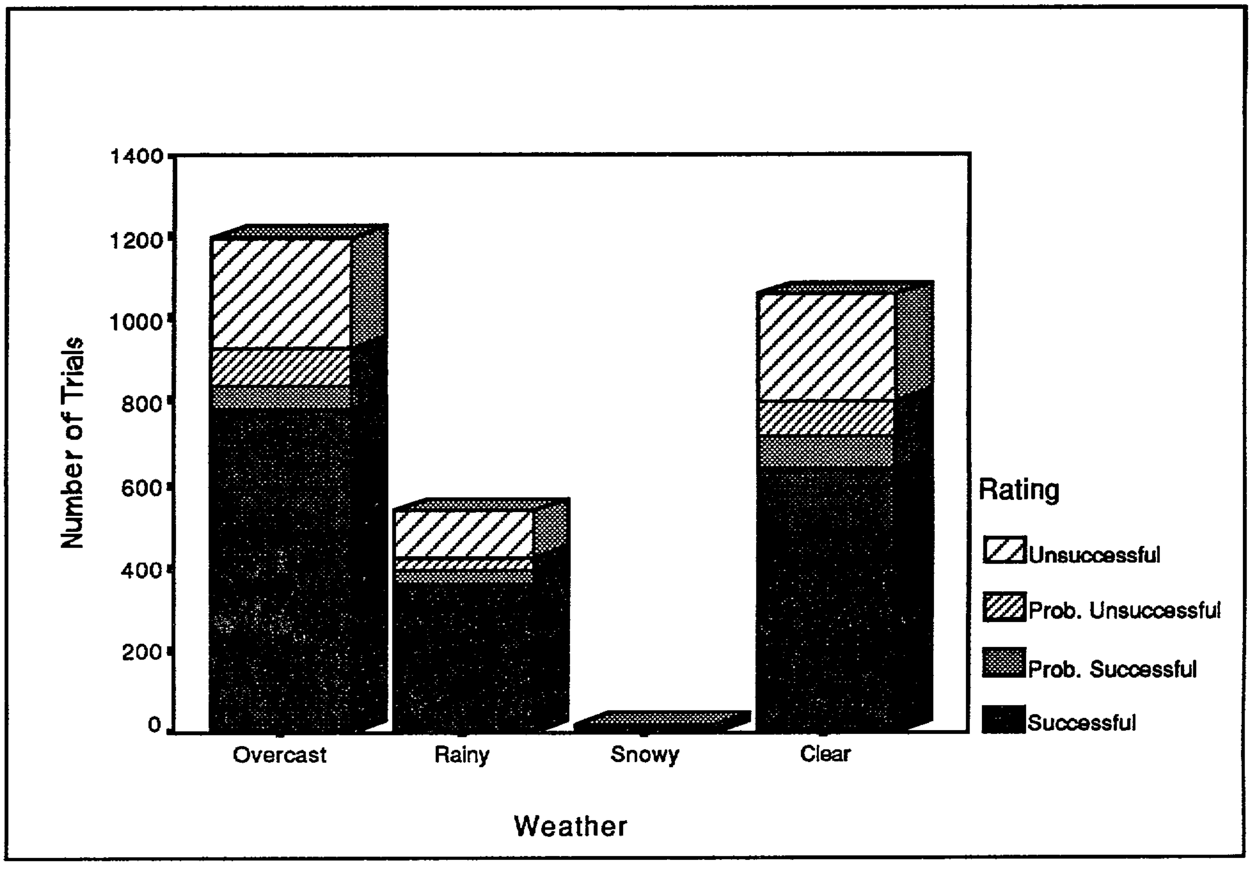


Figure 2.25 - XYPOINT Trial Success by Weather (Number)

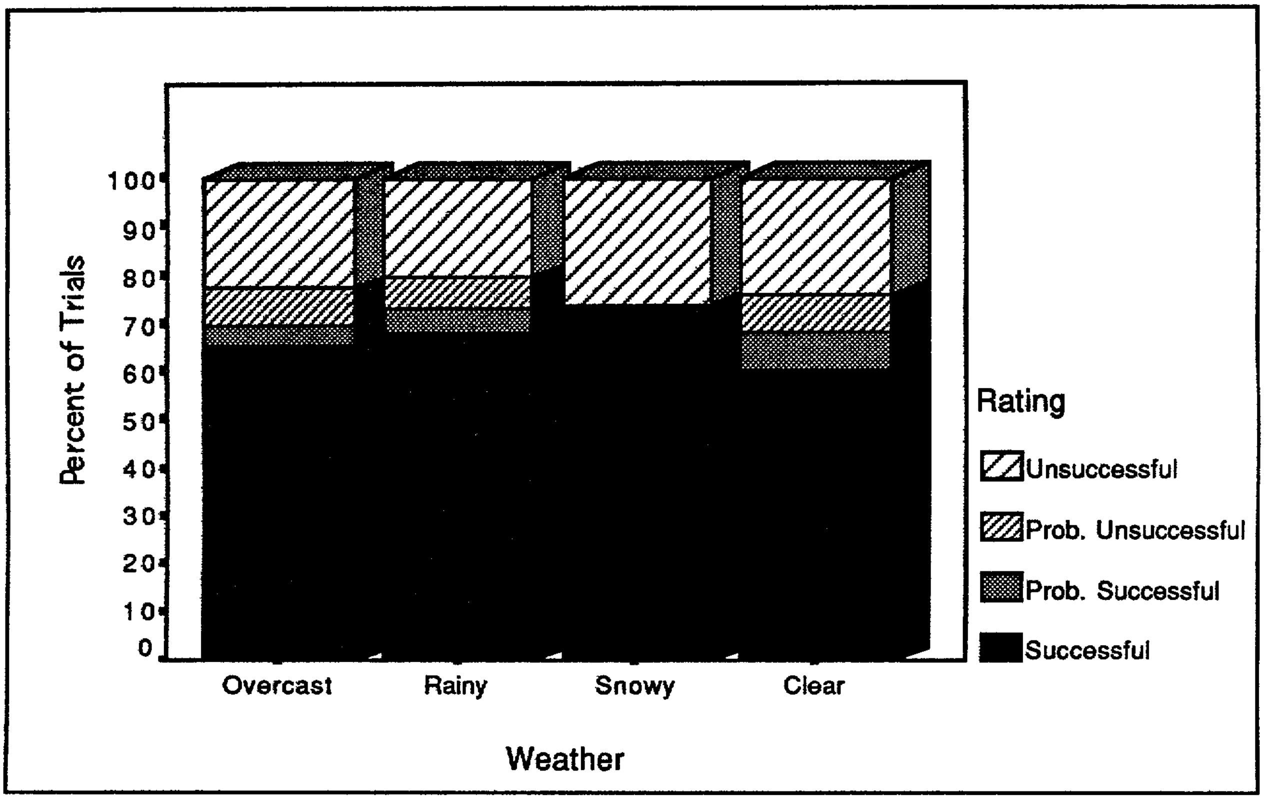


Figure 2.26 - XYPOINT Trial Success by Weather (%)

2.2.6.7 XYPOINT - Relationship Between Trial Success and User Reported Trial Location

Figures 2.27 and 2.28 present percentage and number of successful trials by user reported locations. Figure 2.27 presents the *number* of successful and unsuccessful trials by location, while Figure 2.28 presents the *percent* of successful and unsuccessful trials by location. A Kruskal-Wallis one-way analysis of variance of ratings by location revealed a statistically significant difference in ratings by location (H= 13.54; df = 2; p < .01). The system performed best in urban areas, nearly as well in suburban areas, and and showed some degradation in rural areas.

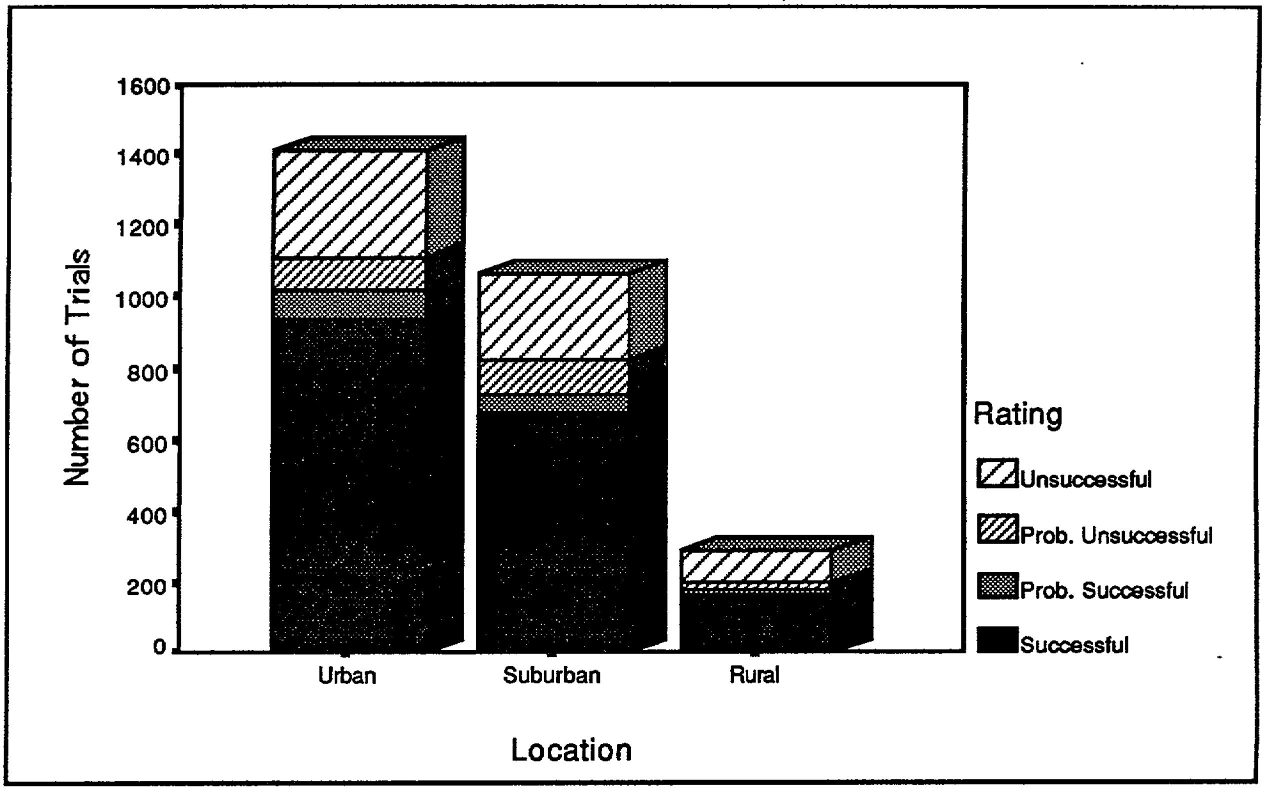


Figure 2.27 - XYPOINT Trial Success by Location (Number)

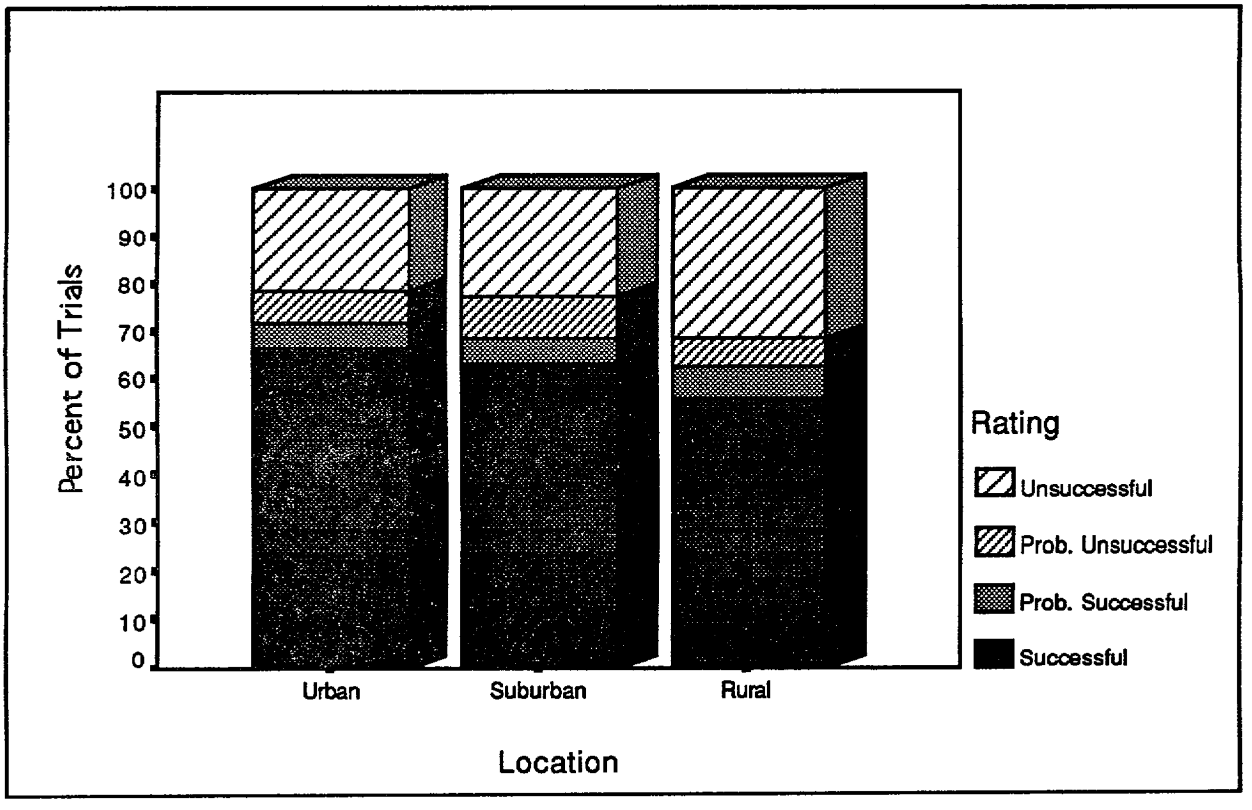


Figure 2.28 - XYPOINT Trial Success by Location (%)

2.2.6.8 XYPOINT - Relationship Between Trial Success and Individual In-Vehicle Unit

Figures 2.29 and 2.30 are extremely difficult to read because of the number of units involved in trials, but they present the results of trial ratings by in-vehicle unit. A Kruskal-Wallis one-way analysis of variance of ratings by unit revealed that there was a statistically significant difference in ratings by unit (H = 726.97; df = 111; $p < .0001$). Some units performed better than others and, in general, the units that did not perform well took part in fewer trials.

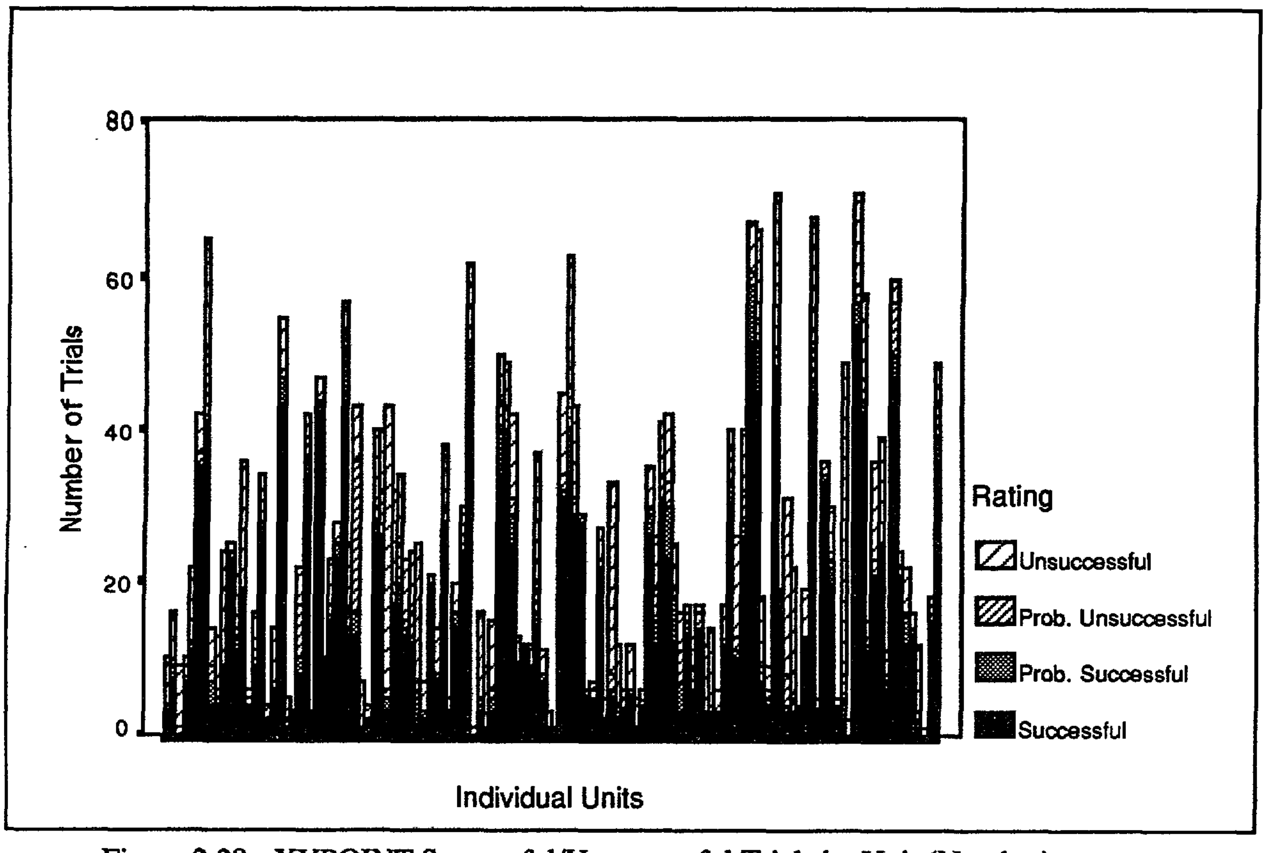


Figure 2.29 - XYPOINT Successful/Unsuccessful Trials by Unit (Number)

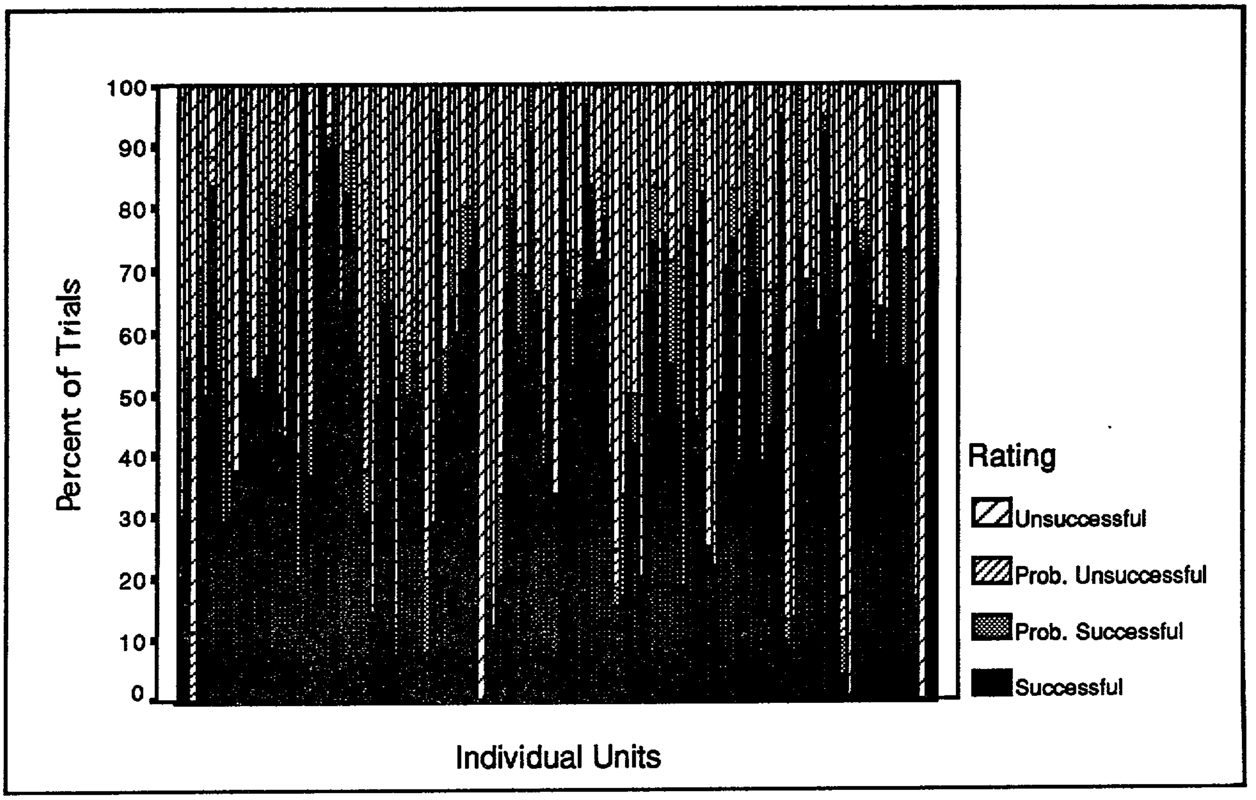


Figure 2.30 - XYPOINT Successful/Unsuccessful Trials by Unit (%)

2.2.7 CONCLUSIONS AND DISCUSSION

The user group deployment test produced large amounts of performance data. While nearly half of all scheduled trials either never occurred or produced unusable results (most never occurred), over 5,000 trials conducted between November 1995 and May 1996 produced analyzable data, an average of more than 700 a month. Nearly two thirds of these (3,129) were trials of the XYPOINT system.

2.2.7.1 Success Rate of Systems

It is difficult to say precisely what is an acceptable level of success for connection and communication of an emergency response call. One could easily argue that nothing short of 100% will do. From this perspective, the 70.5% success rate of all trials (88% successful or probably successful for Motorola, 66% successful or probably successful for XYPOINT) could be viewed as somewhat disappointing. There are, however, numerous other factors to consider.

For this test, volunteers were instructed to push the button only once and to indicate a failure if this single button push failed to result in a connected call. In actual use, a driver would likely reset and push the button again if a successful connection wasn't achieved the first time. Assuming each button push to be independent, an 88% success rate for one trial would increase to 98.5% for two successive trials and over 99% for three; a 66% success rate for one trial would increase to 88.5% for two successive trials and to 96% for three. Of course a second or third button push might not always be possible and would increase response time as measured from the initial button push.

Another set of factors to consider is related to the incomplete development stages of devices and the impermanent infrastructure and personnel. Devices were not market ready, nor were they maintained at the level of a product. The communication infrastructure, especially in the XYPOINT case, was still in development. For example two months after the test, XYPOINT's CDPD communications were changed from piggy backing on existing cellular communications to the use of dedicated lines. Another factor was human error; CSC operators were handling a fairly large cognitive load with minimal training and a test could be classified as unsuccessful if the operator failed to properly enter verification.

If we assume:

(1) multiple button pushes,
(2) improvements in hardware and support as required by the marketplace,
(3) professional operators,
(4) a more universally deployed cellular infrastructure, and
(5) the absence of obvious topographic interference,

then based upon the results of the user deployment test it is reasonable to believe that from the strict perspective of achieving successful connection and communication, these systems could approach 100% successful operation in the near future.

2.2.7.2 Response Time

For this discussion, "response time" means the time it took for the emergency operator to answer and exchange initial information about the call, not the time for service to be provided. As with success, it is difficult to precisely define an acceptable length of time between initiating an emergency call and exchanging information with an operator. Not only do response requirements differ with the type of emergency, but response time will also be impacted by demand, that is, how widespread the emergency is (e.g. car accident vs. earthquake). While in the user deployment test volunteers pushed different buttons representing different types of emergencies, trials were conducted independent of emergency type (i.e. all calls were handled identically) and demand (i.e. no effort was made to measure the impact of simultaneous calls on response time).

Unfortunately, while there is available data on the time spent by 911 emergency operators in *handling* a call (see DEA's *Institutional Issues Report),* no data could be found on the time spent by 911 callers ***waiting*** for their calls to be answered. In any case, this is more a political and economic issue (how many operators for how many people?) than a technical one. It may be worth noting that a search of articles on problems with 911 systems revealed a focus on issues relating to dispatch and how calls were handled rather than on time spent in connecting.

Even without these guides to an acceptable "response time," it is clear that in the user deployment test, calls were connected and initial information was exchanged quite rapidly. From the users perspective, of the 3,232 trials for which we have response time data, 2,287 (over 70%) were reported as being verified within two minutes, and only 112 (3.5%) were reported as taking longer than five minutes. From the CSC data, of the 3,038 trials with usable time information, 2884 (95%) were verified within two minutes and only 24 of 3,038 (less than 1%) took more than five minutes. The much stronger CSC results are likely due to the fact that the machine time began with call receipt, while user perceived time began with the button push. In addition, users tended to round time measurements off to the minute.

Overall, the user deployment test indicates that with respect to the time required to connect to and exchange initial information with emergency operators, the short connect and response times demonstrated have the potential to facilitate emergency response to on-road incidents. If this contributed to a decreased incident response time, then reduced congestion and associated benefits would result.

2.2 7.3 Other Conclusions

A number of other potentially useful conclusions can be drawn from the user deployment test.

The success by month analysis showed an improvement in performance of both systems over the initial course of the operational test. For both systems, during the first several months of testing some of the "kinks" were ironed out and performance was best near the end of the operational test. In both cases, however, the final month of testing was not the most successful, indicating a leveling off below perfect performance once implementation "bugs" were eliminated.

The success by time of day analysis did not conclusively support our hypothesis that performance would be impaired during times of peak cellular use, presumably the commute and **lunch hours when people were in their cars. Unfortunately the Motorola data covered only the afternoon from 2 PM to 6 PM, and if anything performance tended to improve rather than degrade as** the **afternoon progressed. The XYPOINT data, however, did show its worst** performance during heavy cellular periods (7 - 8 AM and 5 - 6PM), indicating that the CDPD network was adversely affected by high cellular usage or operator overload. Due in part to these results, AT&T came to the decision to begin dedicating some cellular lines to CDPD.

Both systems experienced relatively low success rates during the earliest periods of the day **(Motorola's lowest success rate occurred during the first hour of CSC operation, and** XYPOINT's second lowest success rate occurred during the first hour of CSC operation). One explanation for this pattern is that the CSC computer had to be restarted and certain procedures executed each day prior to testing. This startup process could account for the fact that for each system, some of the worst performance occured during the first hour of operation.

Finally, analysis of success by individual unit indicates that test results were not skewed by the malfunctioning of a few units. When units consistently malfunctioned, they were replaced before being used in large numbers of trials.

This concludes the discussion of the user deployment test. The following section discusses the performance test we called "Simulated Service Delivery."

2.3 SIMULATED SERVICE DELIVERY TEST

While the PuSHMe evaluation did not involve actual emergencies, the Simulated Service Delivery Test incorporated simulated dispatch and delivery of emergency services. These trials covered aspects of the PuSHMe systems from call initiation to simulated service arrival and closure. (See Figure 2.1 on page 12 for an overview of the test procedure.)

2.3.1 OBJECTIVES OF TEST

The objectives of the simulated service delivery test were to evaluate (1) how long it took from the time a volunteer pressed a button to the time service was dispatched, and (2) the quality of the PuSHMe information provided to simulated service providers. Because these tests were simulations of responses to emergency calls, they were not used to measure the time it took an emergency service provider to actually reach a volunteer. However since service providers did travel to the simulated incident, some measure of PuSHMe's potential impact on travel time to the incident scene was obtained by asking these service providers to compare the quality of the location information to that usually available from cellular 911 calls.

2.3.2 TEST METHODOLOGY

The simulated service delivery test was conducted between November 1995 and May 1996 using both Motorola and XYPOINT devices. For this test, six DEA or WSDOT employees (not participants in the User Group Deployment Test) were provided with user logs which identified a location they were to go to and an emergency scenario that they were to "act out." These volunteers drove to the location and pressed the button appropriate to the type of emergency simulated. The operator at the CSC: (1) received the call, (2) entered in the trial number, (3) confirmed the nature of the call, (4) confirmed the location of the caller relative to the location automatically displayed on the desktop map, (5) asked the volunteer a series of questions relating to the simulated incident, and (6) called a dispatch center, asking them to send an emergency service professional to the vehicle based on the PuSHMe information. Participating service providers came from the King County Police, the American Automobile Association (AAA), and WSDOT incident response.

The key difference between the Motorola and XYPOINT trials was how the incident infoxmation was exchanged between the users and the CSC operators. As with the user deployment test, the Motorola users verbally exchange-d information with the CSC operators, while the XYPOINT users read messages on a screen display in their vehicle and responded to messages with "Yes" or "No." In each case, after the CSC operator called a dispatch center asking them to send out assistance, the user waited and notified the CSC of the arrival of assistance. (See Appendix D for instructions provided to users and operators.)

Once the CSC operators had: (1) confirmed the nature of the request, (2) located the vehicle on their on-screen map, and (3) confirmed this location with the volunteer, the CSC operators then called the service dispatcher at King County Police or AAA. The CSC operators provided the dispatcher with information about the simulated incident (including the location), and the dispatchers then relayed this information to their service providers who actually drove out to the simulated incidents.

All of the dispatchers and service providers involved in the test were asked to complete a brief survey regarding the PuSHMe trials. Because the simulated service providers and dispatchers had other "real" emergencies or road service calls to attend to, the PuSHMe team wanted to minimize the time required to complete the survey. Therefore, the survey was intentionally kept very brief and simple to complete. Service providers and dispatchers were asked the following four questions:

Question 1. How useful was the information describing the incident (other than location)?

Question 2. How accurate was the location information?

Question 3. How did the information describing the incident (other than location information) compare to information typically provided from cellular 9 11 calls?

Question 4. How did the location information for this call compare to location information typically provided from cellular 911 calls?

Service providers and dispatchers were also asked if there was other information that would be helpful for them to receive from a system like PuSHMe to help them better respond to on-road emergency calls. There was also space provided on the bottom of the survey for additional comments.

2.3.3 DATA ANALYSIS

For the simulated service delivery trials, data was obtained from the in-vehicle volunteers, the CSC operators/computers, the simulated service dispatchers, and the simulated service providers.

Data Obtained from In-vehicle Volunteers

1. Trial number
2. Date
3. Location
4. Scenario
5. Button pushed
6. Time button pushed
7. Time trial confirmed by operator (XYPOINT)/call answered by operator (Motorola)
8. Time operator notified service provider
9. Time operator indicated service was dispatched

Data Obtained from CSC Operators/Computers

1. Test type and number
2. Time service notified
3. Time service dispatched

Data Obtained from Simulated Service Dispatchers

1. Quality of location information
2. Quality of incident information

Data Obtained from Simulated Service Providers

1. Quality of location information
2. Quality of incident information

The volunteer report forms and CSC data logs were linked to evaluate the simulated service delivery tests. Data was analyzed to determine the time required to dispatch simulated service, and the simulated service provider response logs were evaluated to determine the quality of the location and incident information.

For this test, in-vehicle volunteers recorded time using Seiko™ watches, which keep time based on GPS. However, the CSC computer clocks did not consistently reset to GPS time. Thus while having a similar time reference helped in linking CSC data to user report form data, the inconsistent time difference between the Seiko™ watches and CSC clocks prevented us from combining CSC and volunteer times in a single calculation. Instead, volunteer reported time differences and CSC time differences were calculated separately.

2.3.4 RESULTS: TIME TO DISPATCH SERVICE

Because the Motorola and XYPOINT systems generated slightly different data, results are presented separately.

2.3.4.1 Motorola Time to Dispatch Service

As shown in Table 2.3, the CSC data indicates that once the call was received at the CSC, it took an average of about four minutes to notify the service provider; and an average of an additional one and three quarters minutes to: (1) receive confirmation that the service provider was dispatching service, (2) indicate this information to the user, and (3) indicate in the data field that service was dispatched. The total average time it took, from receipt of call at the CSC, through the process of service dispatch was about six minutes.

The results indicated on the user logs were similar. Volunteers indicated that it took an average of just over a minute for the call to be answered after the user pressed the button, and once answered it took about four and a half minutes for the operator to indicate to the user that service was dispatched. Based on user log data, the total average time it took from button push to service dispatch was about five and a half minutes.

Table 2.3 - Motorola Simulated Service Delivery Test Results

	CSC Data			**User Log Data**				
	Call Receipt To Svc. Notif.	Svc. Notif. To Svc. Dispatch	Call Receipt to Svc. Dispatch	Button Push to Call Answer	Button Push to Svc. Notif.	Svc. Notif. To Svc. Dispatch	Call Answer to Svc. Dispatch	Button Push to Svc. Dispatch
Avg.	0:03:55	0:01:43	0:05:55	0:01:05	0:03:28	0:02:06	0:04:25	0:05:31
St. Dev.	0:01:58	0:03:33	0:04:10	0:01:10	0:02:17	0:01:09	0:02:15	0:02:24
Min.	0:01:30	0:00:10	0:02:13	0:00:00	0:01:00	0:00:30	0:01:00	0:02:00
Max.	0:10:54	0:12:27	0:29:47	0:09:32	0:14:10	0:05:13	0:14:00	0:14:40
N.	37	29	63	79	78	77	78	78

A few items in Table 2.3 require further explanation. Note that the Motorola User Log Data indicating average time from button push to service dispatch (0:05:31) to be slightly less than the CSC Data indicating the average time from call receipt to service dispatch (0:05:55). These results are counterintuitive since the User Log time includes over a minute of wait time between button push and call receipt. However, the CSC Data time includes time spent by operators between verbally reporting to the user that service was dispatched and actually entering in the data log that service had been dispatched. Perhaps operators talked to the user about the trial for a short time or waited until after closing the call to indicate in the data that service was dispatched for the trial. In any case, operator comments indicated that it was not unusual for them to enter the data record indicating service dispatch a minute or two after informing the volunteer that service had been dispatched Similarly, the unusually long maximum time in the CSC data (more than twice that indicated in the user log) probably was due to a long delay in the operator entering the service dispatch data record.

The sample sixes also require further explanation. Note that the sample sizes in the CSC Data for the "Call Receipt to Service Notification" (n = 37) and "Service Notication to Service Dispatch' (n = 29) fields are quite a bit smaller than the other fields. This is because the operators did not always indicate in the data when service was notified. "Service Notified" was meant to represent when the operator informed the user that they had all the information they needed and were now able to notify the service provider. "Service Dispatch" was meant to represent when the operator received an indication from the dispatcher that assistance75s
failed to enter either or both codes; sometimes they entered in a code indicating that service was notified and dispatched at the same time (perhaps they were not aware of the difference between the two actions). For cases where operators indicated that service was notified and dispatched at the same time, the time stamp was taken to represent service dispatch only.

2.3.4.2 XYPOINT Time to Service Dispatch

As shown in Table 2.4, the XYPOINT CSC data indicates that once the call was received at the CSC, it took an average of about nine minutes to notify the service provider and approximately an additional one and one quarter minutes to (1) receive confirmation that the service provider was dispatching service and (2) indicate in the data field that service was dispatched. The total average time it took from call receipt to dispatch service was about ten and three quarter minutes.

The results indicated on the user logs were similar. Volunteers indicated that it took slightly less than a minute from button push to confirmation of the request by the operator, and once confirmed it took about eleven minutes for the operator to indicate to the user that service was dispatched Based on user log data, the total average time it took for the operator to dispatch Service after the user pushed the button was about eleven and three quarters minutes.

Table 2.4 XYPOINT Simulated Service Delivery Test Results

	CSC Data			User Log Data		
	Cdl Received to Service Notified	Service Notified to Service Dispatched	Call Received to Service Dispatched	Button Pushed to Button Pushed Confirmed	Button Push Confirmed to Service Dispatched	Button Pushed lo Service Dispatched
Avg.	0:09:01	0:01:16	0:10:41	0:00:57	0:10:58	0:11:46
St. Dev.	0:05:39	0:01:14	0:06:00	0:01:04	0:06:15	0:06:20
Min.	0:02:37	0:00:04	0:03:14	0:00:00	0:03:30	0:04:00
Max.	0:29:26	0:04:49	0:32:12	0:04:18	0:30:00	0:31:00
n	50	40	40	45	42	45

Note that for both the XYPOINT and Motorola systems there were more user log trials than CSC data trials. This is because it was not possible to link every user log trial to CSC data. In the CSC data, most trials were identified by trial number. Where trial numbers in the user log data were not found in the CSC data, the evaluators tried to identify the trials even without the trial number (by identifying simulated service delivery trials without a trial number conducted by the same vehicle at the same time). Still, not every trial indicated on user report forms could be linked to CSC data. Possible explanations include: (1) the CSC computer was not recording data at the time, or (2) some erroneous information was on the user report form.

23.5 CONCLUSIONS: TIME TO SERVICE DISPATCH

For the purposes of this test, the primary difference between the systems was the method of receiving the incident information from the in-vehicle user: the Motorola system enabled CSC operators to verbally ask the user a series of questions before contacting the simulated service provider, while the XYPOINT system required the operator to send a series of questions (in text format) and wait for responses before contacting the simulated service providers. Not surprisingly it took nearly twice as long for a text-based system to exchange pre-dispatch information between volunteer and CSC operator than a voice-based system. The impact of an additional six minutes on the effectiveness of an emergency response system is another question.

Under the PuSHMe scenario, both the XYPOINT and Motorola systems can be seen as adding additional time to the emergency response process as compared to a typical cellular call to an E-911 center. The CSC operators first answered calls from the users and gathered incident information from them (either via voice or text), and then called the simulated emergency service providers and passed this information on. The Motorola operators took an average of 103 seconds to transmit this information to the dispatcher while the XYPOINT operators took an average of 76 seconds to transmit this information to the dispatcher. As reported by David Evans and Associates (***Institutional Issues Report,*** 1996), for cellular calls observed at E-9 11 centers, the average call duration for au incident call was 144 seconds for the Washington State Patrol in Bellevue, and 105 seconds for the King County Public Service Answering Point (PSAP). In other words, the information exchange between user and CSC operator did not significantly shorten the time it took a PuSHMe CSC operators to communicate the incident information to service providers as compared to the time it takes a user to communicate this information directly.

Of course time is not the only issue. The pre-screening of calls may have other benefits that outweigh the possible time lost. One of these benefits, ‘the possible elimination of unnecessary calls to E-91 1 centers, is discussed under institutional issues (Chapter V). Another possible benefit is better location information. If, for example, an in-vehicle user is unable to communicate his or her location to a PSAP, then an additional step that automatically provides location information could reduce the time required to provide emergency services. Trials simulating these conditions (“blind trials”) were scheduled, but the simulated service providers were unable to participate in all of the scheduled trials since they had real emergencies to attend to and therefore had to cancel trials. Ultimately, “blind trials” were not completed. Under “blind trial” circumstances, even with the additional delay associated with an extra call between CSC and PSAP, the total time required to determine the location of the victim and dispatch service could very likely be much less than if the location information were unavailable.

Additional discussion of location issues occurs in the following section and in Chapter V (Institutional Issues).

2.3.6 RESULTS: QUALITY OF INFORMATION

Both dispatchers and service providers were surveyed as to the quality of information provided by PuSHMe response center operators. The dispatchers handled calls from both the XYPOINT and Motorola systems without regard to the specific system that was being tested. In addition, because they were busy fitting the PuSHMe calls in with actual emergency calls, they tended not to complete a survey for each trial they handled. Instead, most completed the survey after handling calls for several trials. Therefore, dispatcher results cannot be broken down by system and their comments are actually a combined impression of the PuSHMe systems together. On the other hand, service providers completed a survey after each service trip. We therefore have separated quality of information results by source: dispatcher results and service provider results.

2.3.6.1 Dispatcher Results

Because AAA and the King County Police are extremely different types of organizations, we have broken down dispatch results by service organization. Particularly in their comparison of PuSHMe incident information to Cellular 911 incident information, AAA and police dispatchers produced different results.

2.3.6.1.1 AAA Dispatcher Results

For Question 1, "How useful was the information describing the incident (other than location)?' responses were received from 20 AAA dispatchers. As shown in Figure 2.3 1 , 10 dispatchers (50%) felt it was *useful;* 9 (45%) felt it was not very *useful;* 1 (5%) felt that the PuSHMe incident information was very *useful;* and none felt it was ***useless*** or indicated that they ***never received incident information.*** A chi-square analysis of responses revealed that there were significant differences in responses ($x^2 = 25.50$; df = 4; $p < .001$), with the majority of respondents finding the incident information *useful* or not very *useful.* Several dispatchers commented on incident information they would have liked to have received but did not, including: (1) the full name and calling number of the person requiring assistance, (2) the type of crime and whether or not it was still in progress, (3) the exact address, and (4) suspect and vehicle descriptions.

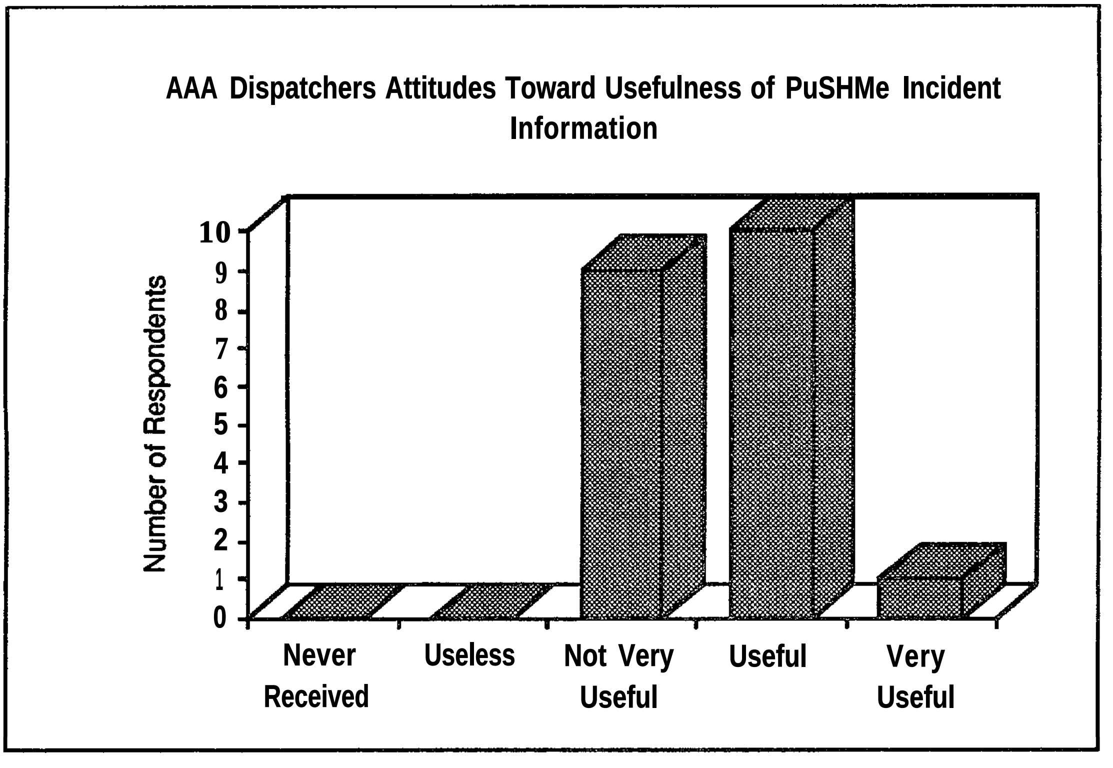


Figure 2.3 1 - AAA Dispatcher Attitudes Towards Usefulness of Incident Information

For Question 2, "How accurate was the location information?' responses were received from 17 AAA dispatchers. As shown in Figure 2.32, 12 AAA (7 1%) dispatchers felt it was ***reasonably accurate;*** 3 (17%) felt it was ***somewhat inaccurate;*** 2 (12%) felt it was very *accurate;* and none felt it ***was quite inaccurate*** or indicated that ***they never received location information*** A chi-square analysis of responses revealed that there were significant differences in responses (x^2 .= 29.18; df = 4; p <.00l), with the majority of respondents finding the location information ***reasonably accurate.***

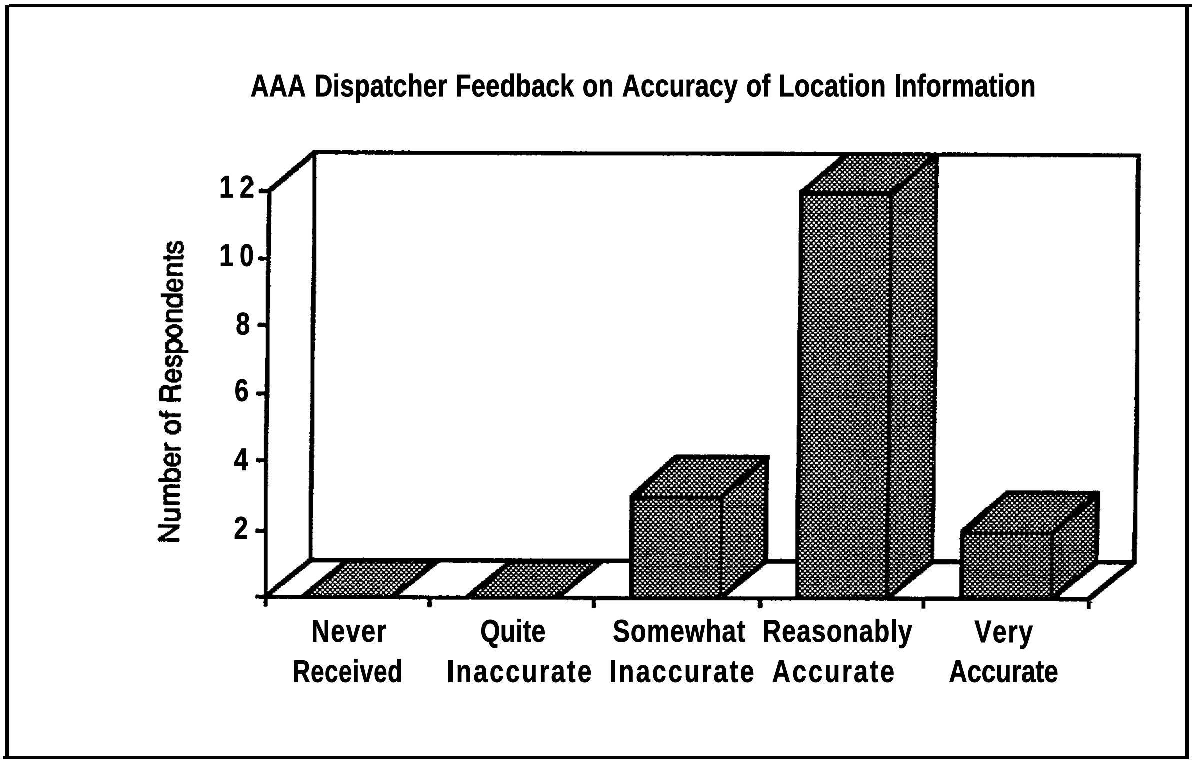


Figure 2.32 - AAA Dispatcher Attitudes Towards Accuracy of Location Information

For Question 3, "How did the information describing the incident (other than location information) compare to information typically provided from cellular 911 calls?" responses were received from 20 AAA dispatchers. As shown in Figure 2.33, 10 AAA dispatchers (50%) felt the incident information was *somewhat* worse; 7 (35%) felt it was about *the same;* 2 (10%) felt it was *much* ***worse;*** 1 (5%) felt it was ***much better;*** and none felt the information was somewhat ***better than*** incident information typically provided from cellular 911 calls. A chi-square analysis of responses revealed that there were significant differences in responses (x^2 = 18.50; df = 4; p $\leq$ .01), with most respondents finding the incident information ***somewhat worse than*** or ***about the same as*** incident information typically provided by cellular 911 calls.

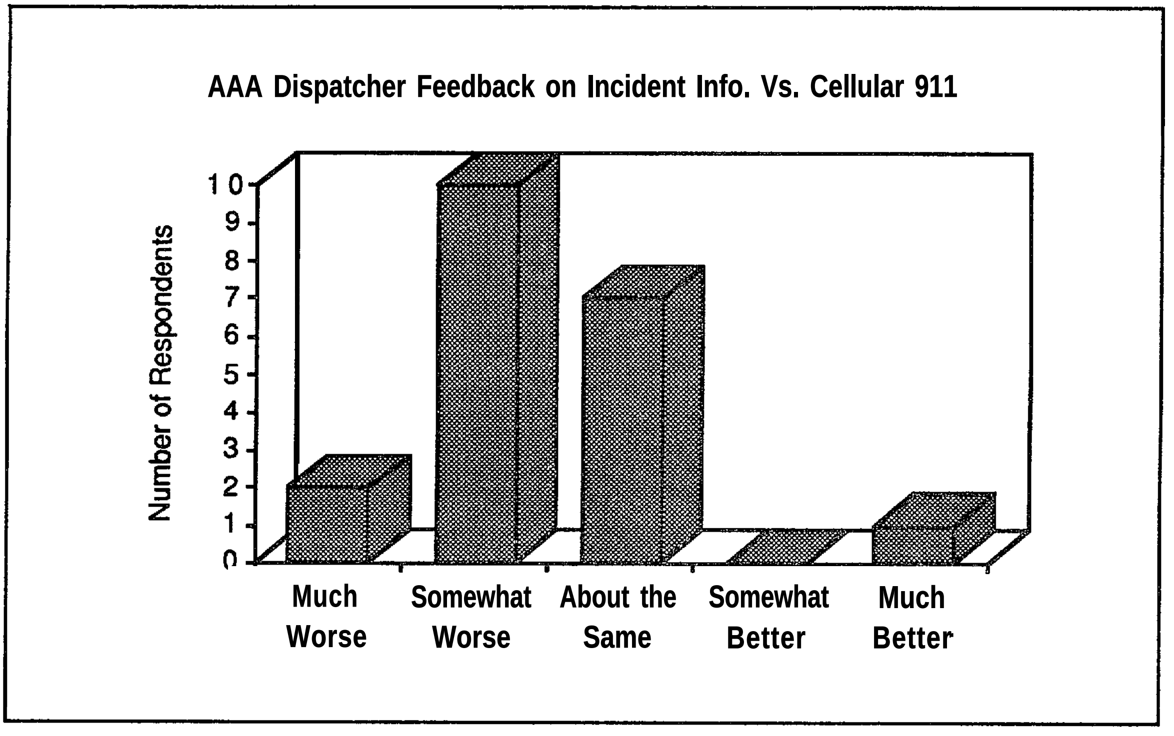


Figure 2.33 - AAA Dispatcher Attitudes Towards PuSHMe Incident Information Compared to Incident Information Received from Cellular 911 Calls

For Question 4, "How did the location information for this call compare to location information typically provided from cellular 911 calls?" responses were received from 20 AAA dispatchers. As shown in Figure 2.34, 10 dispatchers (50%) felt that the PuSHMe location information was *about the same* as the location information typically provided from cellular 911 calls; 5 (25%) felt it was *somewhat better*; 4 (20%) felt it was *somewhat worse*; 1 (5%) felt it was *much better*; and none felt it was *much worse* than typical cellular 911 calls. A chi-square analysis of responses revealed that there were significant differences in responses ($\chi2 = 15.5$; df = 4; $p \leq .01$), with half of respondents finding the location information *about the same* as location information typically provided from cellular 911 calls.

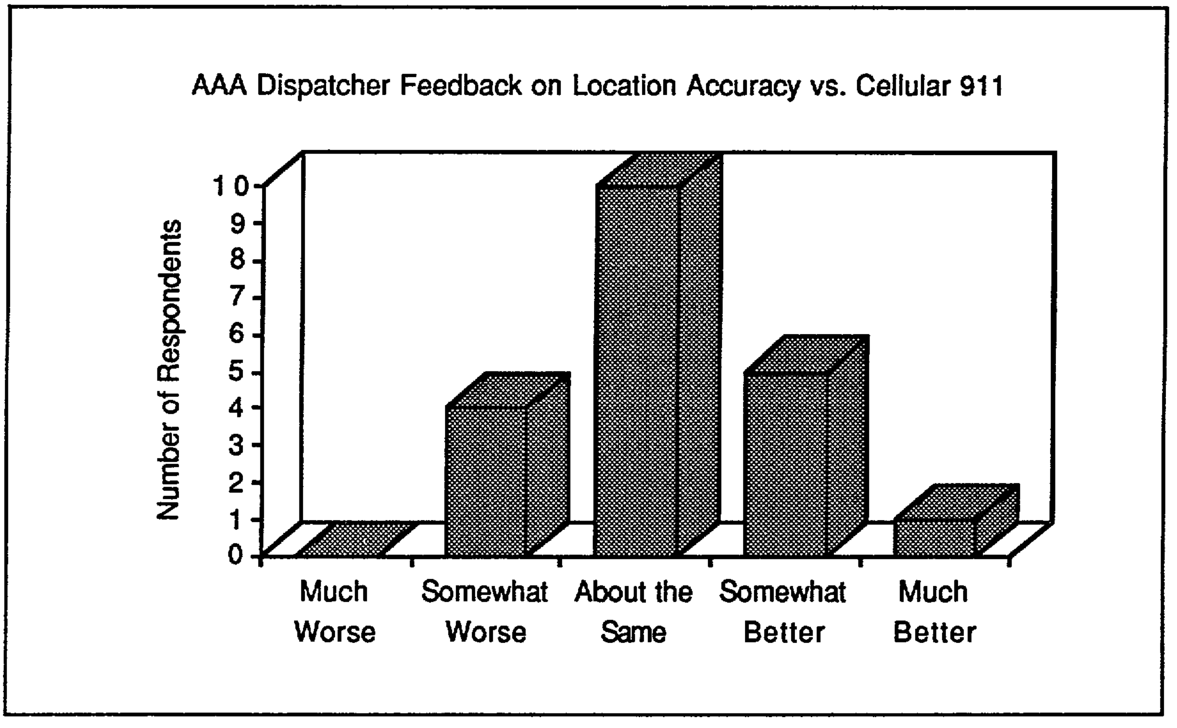


Figure 2.34 - AAA Dispatcher Attitudes Towards PuSHMe Location Information Compared to Location Information Received from Cellular 911 Calls

2.3.6.1.2 King County Police Dispatcher Results

For Question 1, "How useful was the information describing the incident (other than location)?" responses were received from 35 King County Police dispatchers. As shown in Figure 2.35, 20 King County Police dispatchers (57%) felt it was *useful*; 10 (29%) indicated that they *never received incident information;* 5 (14%) felt it was *not very useful*; and none felt it was *very useful* or *useless*. A chi-square analysis of responses revealed that there were significant differences in responses ($\chi 2 = 40.00$; df = 4; $p \leq .001$), with the majority of respondents finding the incident information *useful*. One dispatcher commented that a license plate number and accurate description of the vehicle would have been very useful to receive. (It should be noted that after testing began, it was decided to eliminate specific incident information. When this information was broadcast on the radio, units not directly involved tied up the 'air' with inquiries and requests for clarification (e.g. hearing only part of a 'Robbery' incident.)

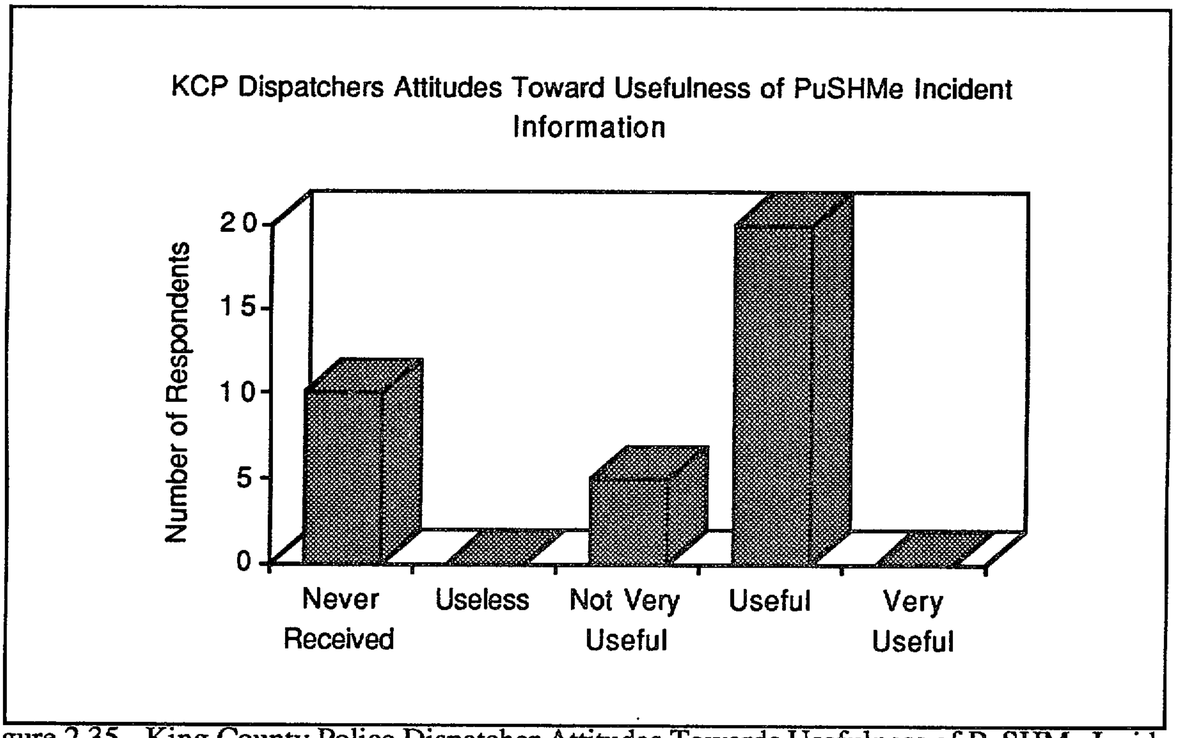


Figure 2.35 - King County Police Dispatcher Attitudes Towards Usefulness of PuSHMe Incident Information

For Question 2, "How accurate was the location information?" responses were received from 35 King County Police dispatchers. As shown in Figure 2.36, 23 King County Police dispatchers ***(66%)*** felt it was *reasonably accurate;* 10 (28%) felt it was very *accurate;* 1 (3%) felt it was *somewhat inaccurate;* 1(3%) felt it was *quite inaccurate* and none indicated that they never ***received location information.*** A chi-square analysis of responses revealed that there were significant differences in responses (x^2 = 55.14; df = 4; p $\leq$ .001), with the majority of respondents finding the location information ***reasonably*** *accurate.* One dispatcher commented that when giving cross streets, an indication of the corner or at least which street the vehicle was actually located on would have been helpful. Another dispatcher commented that it would have been helpful to receive information about the vehicle's direction on the street.

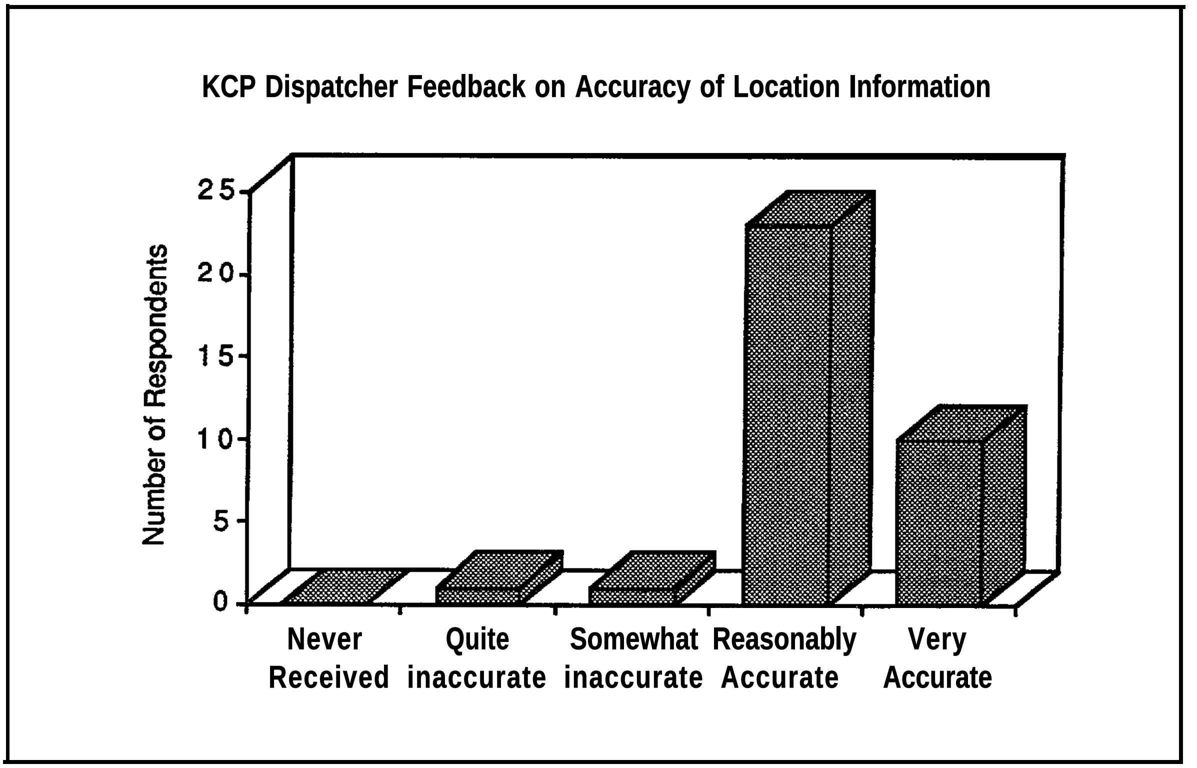


Figure 2.34 - King County Police Dispatcher Attitudes Towards Accuracy of PuSHMe Location Information

For Question 3, "How did the information describing the incident (other than location information) compare to information typically provided from cellular 911 calls?" responses were received from 20 King County Police dispatchers. As shown in Figure 2.37, 10 King County Police dispatchers (50%) felt the incident information was ***about the** same* as information typically provided from cellular 9 11 calls; 7 (35%) felt the information was ***somewhat better***; 1 (5%) felt it was ***much better***; 2 (10%) felt it was much ***worse***; and none felt it was *somewhat* ***worse*** than incident information typically provided from cellular 9 11 calls. A chi-square analysis of responses revealed that there were significant differences in responses (x^2 = 18.5; df = 4; p < .0l), with most respondents finding the incident information ***about the*** same as incident information typically provided by cellular 9 11 calls.

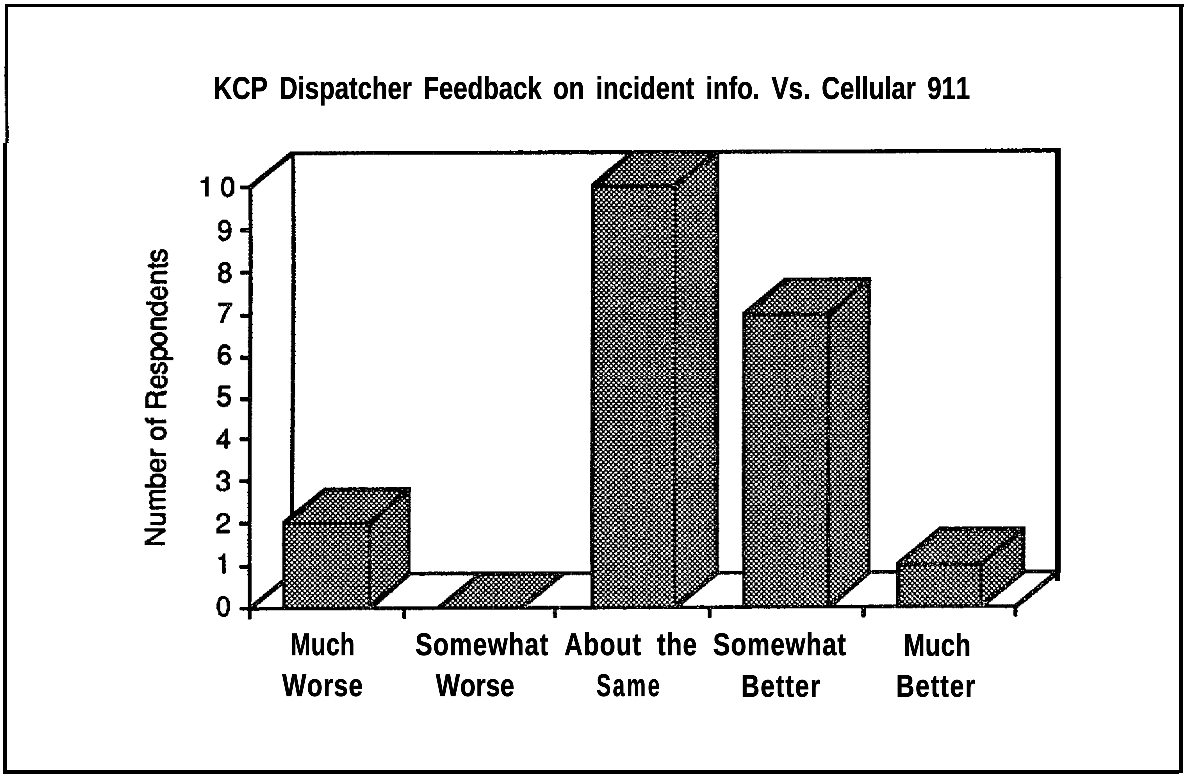


Figure 2.37 - Ring County Police Dispatcher Attitudes Towards PuSHMe Incident Information Compared to Incident Information Received from Cellular 9 11 Calls

For Question 4, "How did the location information for this call compare to location information typically provided from cellular 911 calls?" responses were received from 17 Ring County Police dispatchers. As shown in Figure 2.38, 15 King County dispatchers (88%) felt that the PuSHMe location information was *about the same* as the location information typically provided from cellular 911 calls; 2 (12%) felt it was *somewhat* ***better;*** and none felt it was ***somewhat worse, much better,*** or ***much worse than*** typical cellular 911 calls. A chi-square analysis of responses revealed that there were significant differences in responses (x^2 = 50.35; df = 4; p $\leq$ .001), with most respondents finding the location information ***about the*** *same* as location information typically provided from cellular 911 calls.

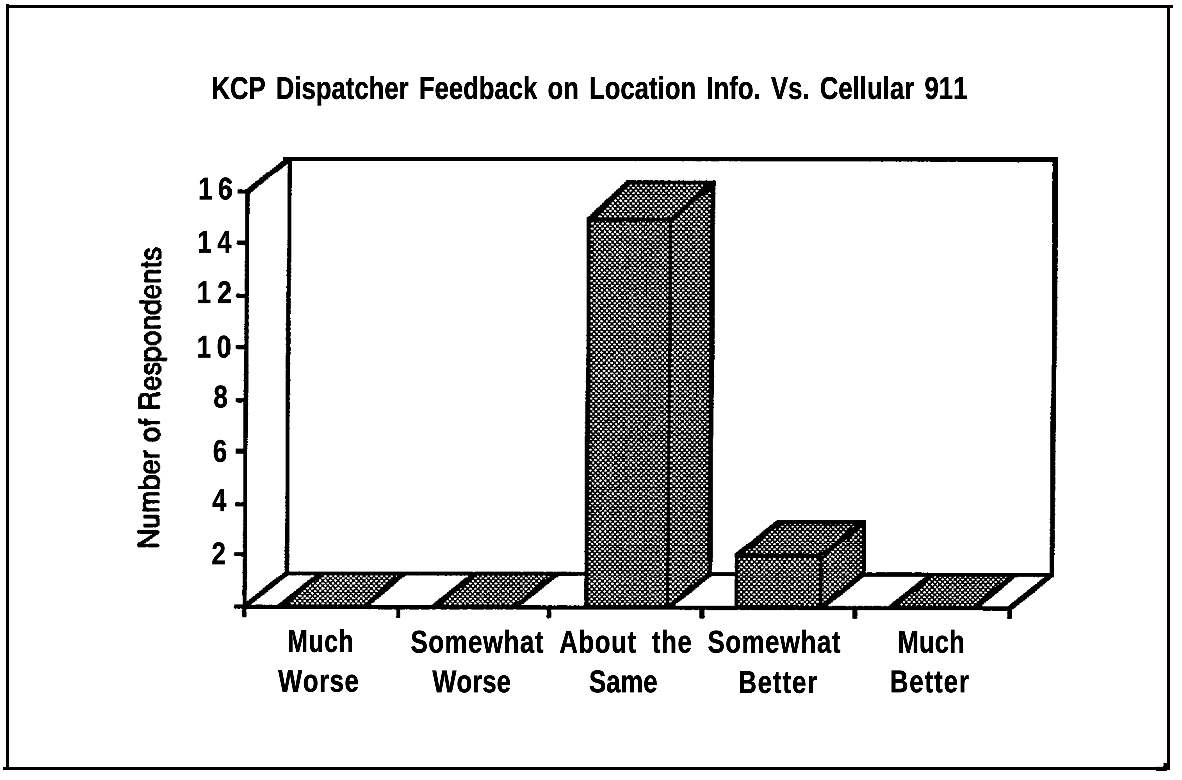


Figure 2.38 - King County Police Dispatcher Attitudes Towards PuSHMe Location Information Compared to Location Information Received from Cellular 911 Calls

2.3.6.2 Service Provider Results

The service providers received information from the dispatchers and drove out to the simulated incident. Unlike the dispatcher results, service provider results can be broken down by system (Motorola versus XYPOINT) since for every trial, the service provider completed a survey upon arrival at the simulated incident. On the other hand, a breakdown of service provider results by service organization would have been difficult to achieve because completed surveys did not consistently identify this information. Therefore, the service provider results are presented by system.

2.3.6.2.1 Motorola Service Provider Quality of Information Results

For Question 1, "How useful was the information describing the incident (other than location)?" responses were received from 74 service providers. As shown in Figure 2.39, 37 service providers (50%) felt it was useful; 20 (27%) felt that the PuSHMe incident information was very *useful;* 9 (12%) felt it was *not very useful;* 8 (11%) indicated that they *never received incident information;* and none felt it was useless. A chi-square analysis of responses revealed that there were significant differences in responses ($x^2 = 51.54$; df = 4; p $\leq .001$), with half of respondents finding the incident information useful.

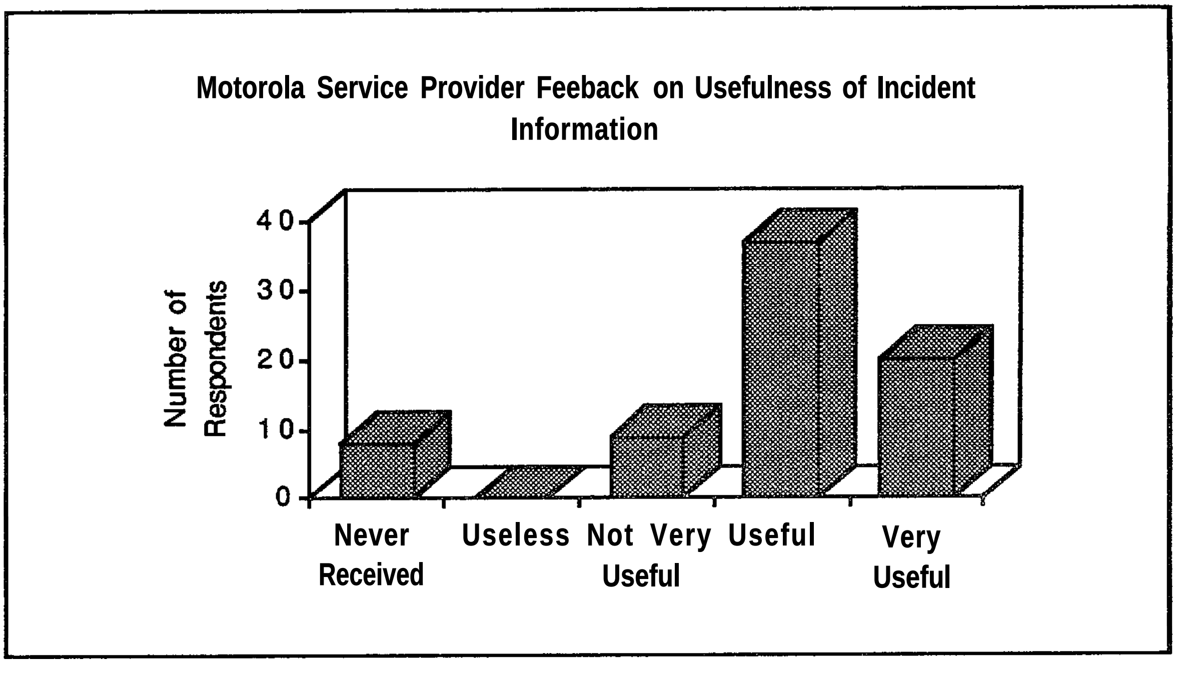


Figure 2.39 - Simulated Service Provider Attitudes Towards the Usefulness of Motorola Incident Information

For Question 2, "How accurate was the location information?" responses were received from 74 service providers. As shown in Figure 2.40, 45 service providers (60%) felt it was very accurate; 21 (29%) felt it was *reasonably accurate*; 5 (7%) felt it was *somewhat* inaccurate; 3 (4%) felt it was *quite inaccurate*; and none indicated that they *never received location information.* A chi-square analysis of responses revealed that there were significant differences in responses (x^2 = 94.92; df = 4; p ≤ .001), with most respondents finding the location information very accurate.

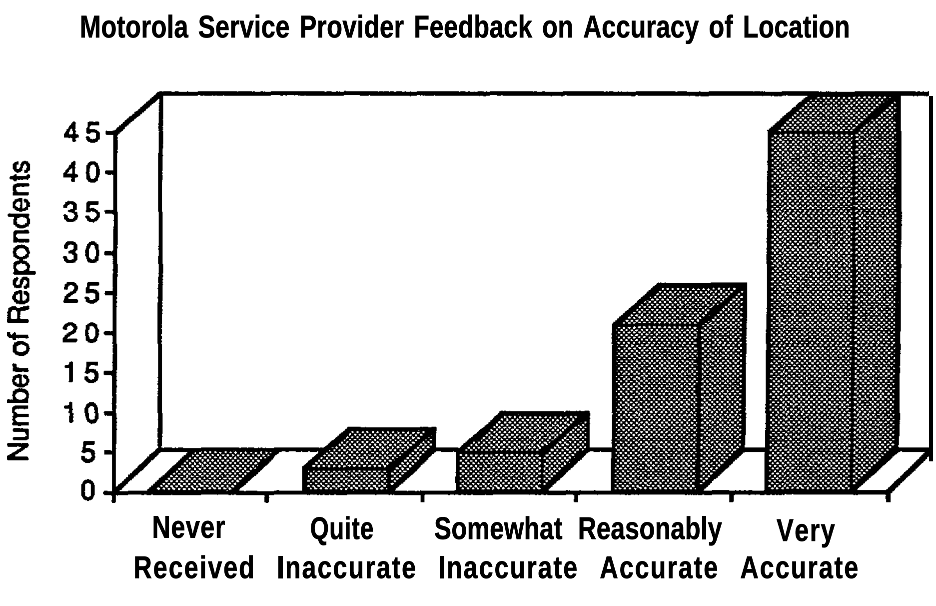


Figure 2.40 - Simulated Service Provider Attitudes Towards the Accuracy of Motorola Location Information

For Question 3, "How did the information describing the incident (other than location information) compare to information typically provided from cellular 911 calls?' responses were received from 66 service providers. As shown in Figure 2.41, 44 service providers (66%) felt the incident information was about *the same* as information typically provided from cellular 911 calls; 8 (12%) felt the information was *somewhat better;* 8 (12%) felt it was *somewhat worse*; 5 (8%) felt it was *much better;* and 1 (2%) felt it was *much worse.* A chi-square analysis of responses revealed that there were significant differences in responses ($x^2 = 92.33$; df = 4; $p \leq .001$), with most respondents finding the incident information *about* the *same* as incident information typically provided by cellular 911 calls.

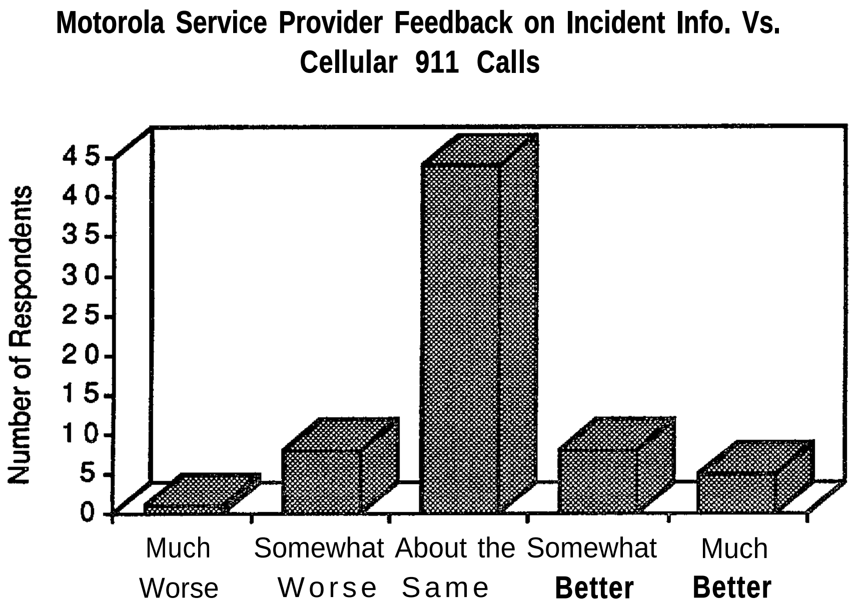


Figure 2.41 - Simulated Service Provider Attitudes Towards the Motorola Incident Information System Compared to Incident Information Received from Cellular 9 11 Calls

For Question 4, ‘How did the location information for this call compare to location information typically provided from cellular 911 calls?’ responses were received from 70 service providers. As shown in Figure 2.42, 46 service providers (66%) felt that the PuSHMe location information was *about the same* as the location information typically provided from cellular 911 calls; 12 (17%) felt it was *much better;* 7 (10%) felt it was *somewhat better;* 4 (6%) felt it was somewhat ***worse;*** and 1 (1%) felt it was much *worse than* typical cellular 911 calls. A chi-square analysis of responses revealed that there were significant differences in responses (χ^2 = 96.14; df = 4; $p \leq .001$), with most respondents finding the location information *about the same* as location information typically provided from cellular 911 calls.

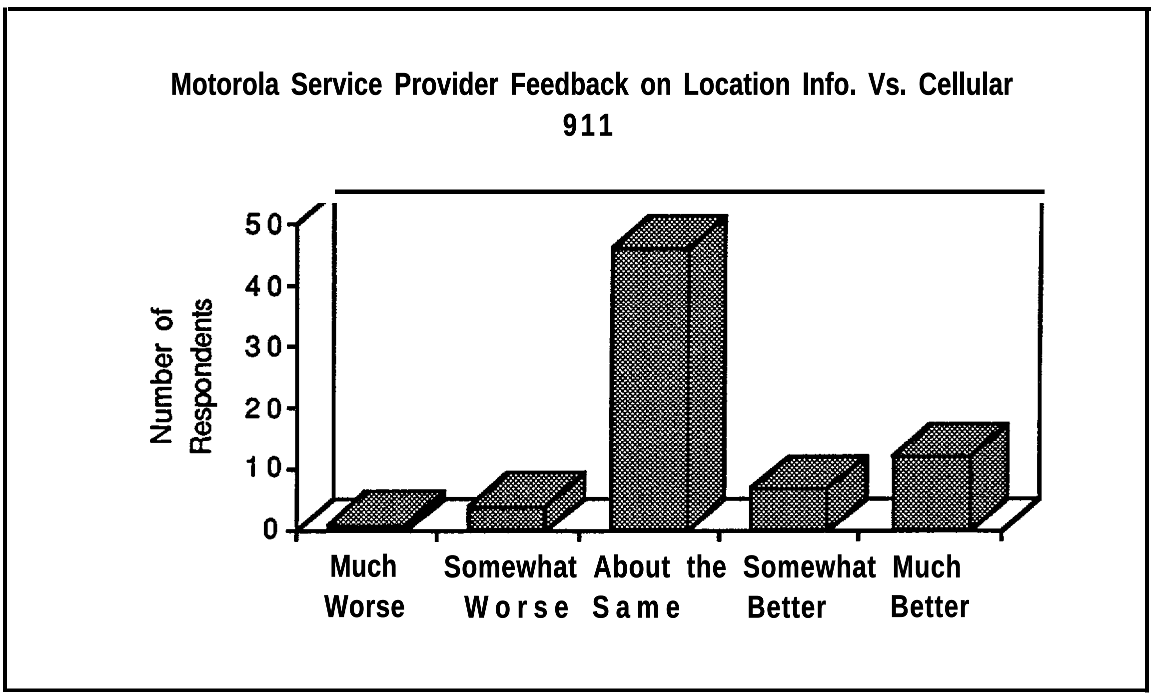


Figure 2.42 - Simulated Service Provider Attitudes Towards Motorola’s Location Information Compared to Information Received from Cellular 911 Calls

2.3.6.2.2 XYPOINT Service Provider Quality of Information Results

For Question 1, "How useful was the information describing the incident (other than location)?" responses were received from 48 service providers. As shown in Figure 2.43, 20 service providers (42%) felt it was *useful*; 14 (30%) felt that the PuSHMe incident information was *very useful*; 6 (12%) indicated that they *never received incident information*; 6 (12%) felt it was *not very useful*; and 2 (4%) felt it was *useless*. A chi-square analysis of responses revealed that there were significant differences in responses ($\chi 2 = 22.21$; df = 4; $p \leq .001$) with the majority of respondents finding the incident information *useful* or *very useful*.

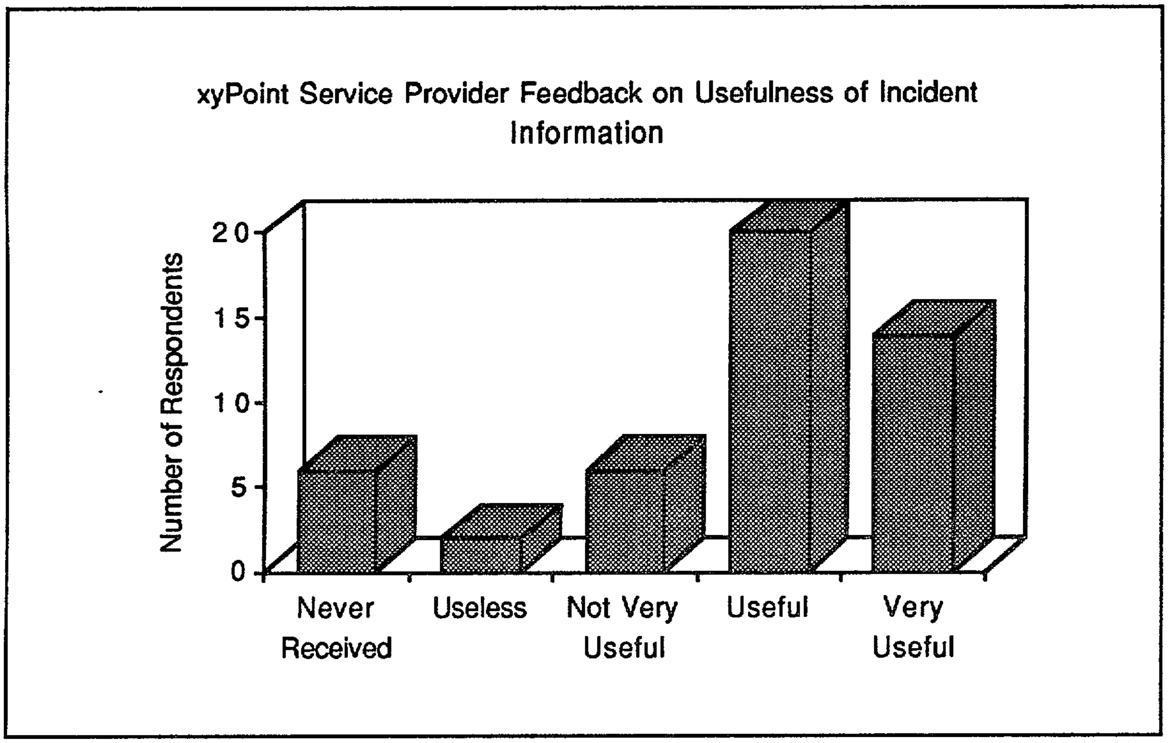


Figure 2.43 - Simulated Service Provider Attitudes Towards the Usefulness of XYPOINT Incident Information

For Question 2, "How accurate was the location information?" responses were received from 48 service providers. As shown in Figure 2.44, 25 service providers (52%) felt it was *very accurate*; 17 (36%) felt it was *reasonably accurate;* 4 (9%) felt it was *somewhat inaccurate*; 2 (4%) felt it was *quite inaccurate*; and none indicated that they *never received location information.* A chi-square analysis of responses revealed that there were significant differences in responses ($\chi 2$ = 49.29; df = 4; $p \leq .001$) with most respondents finding the location description *very accurate* or *reasonably accurate.*

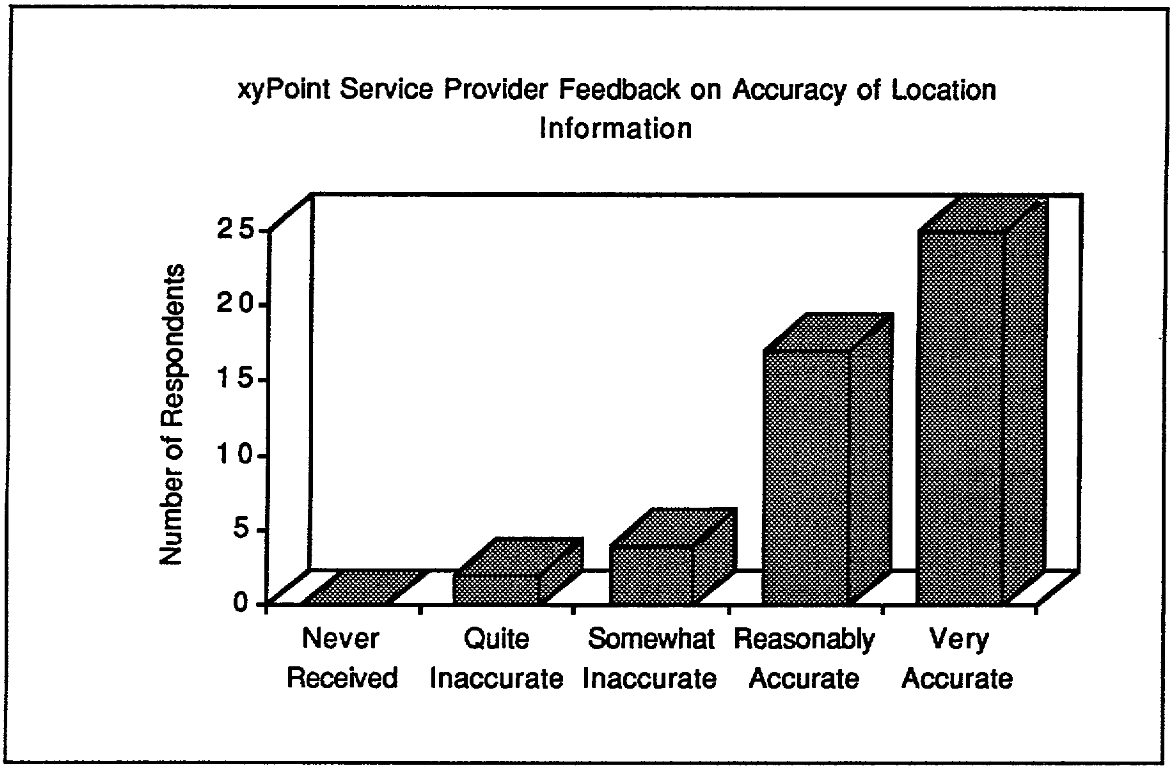


Figure 2.44 - Simulated Service Provider Attitudes Towards the Accuracy of XYPOINT Location Information

For Question 3, "How did the information describing the incident (other than location information) compare to information typically provided from cellular 911 calls?' responses were received from 43 service providers. As shown in Figure 2.45, 25 service providers (58%) felt the incident information was about the *same* as information typically provided from cellular 911 calls; 7 (16%) felt it was *much worse;* 6 (14%) felt the information was *somewhat better;* 4 (9%) felt it was *somewhat worse;* and 1 (2%) felt it was *much better.* A chi-square analysis of responses revealed that there were significant differences in responses ($x^2 = 41.53$; df = 4; $p \leq .001$), with most respondents finding the incident information ***about*** the *same* as incident information typically provided by cellular 911 calls.

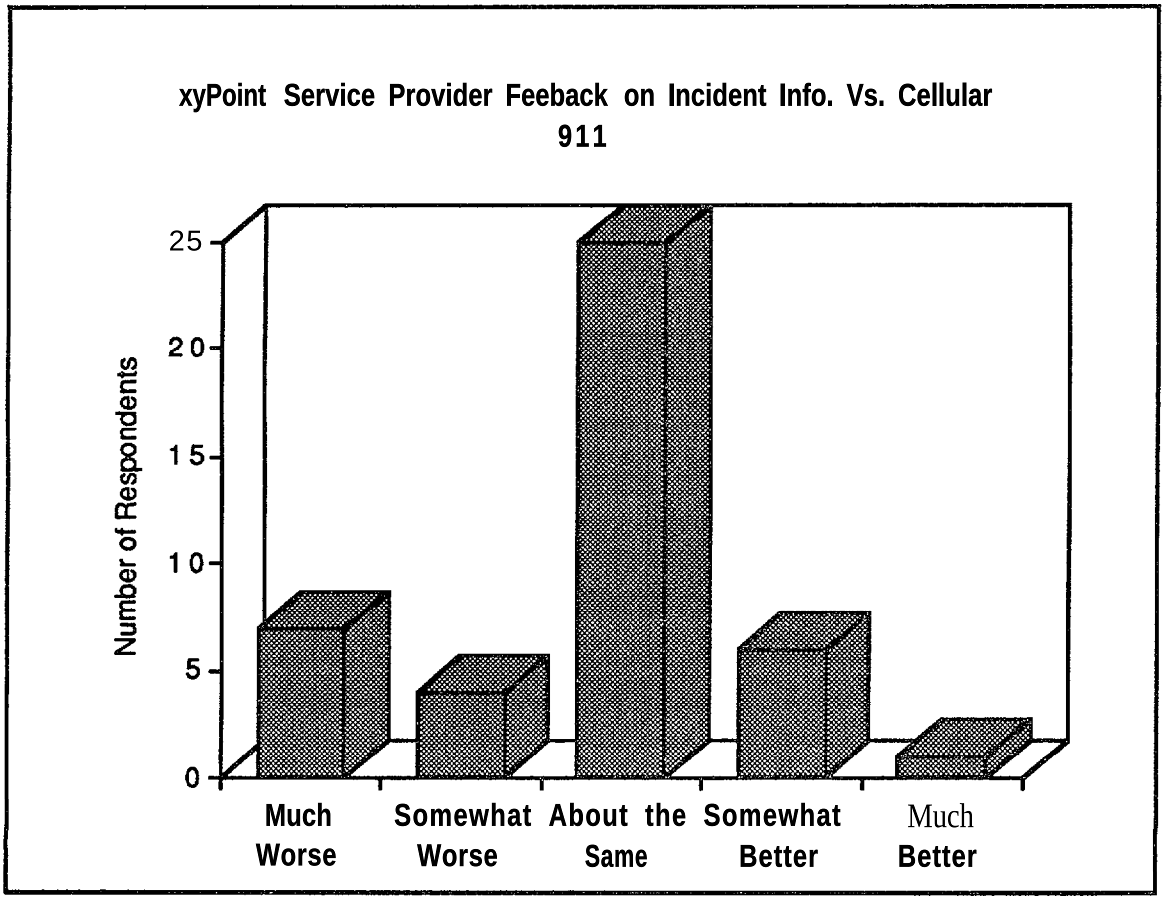


Figure 2.45 - Simulated Service Provider Attitudes Towards XYPOINT's Incident Information Compared to Incident Information Received from Cellular 911 Calls

For Question 4, "How did the location information for this call compare to location information typically provided from cellular 911 calls?" responses were received from 42 service providers. As shown in Figure 2.46, 29 service providers (69%) felt that the PuSHMe location information was *about the same* as the location information typically provided from cellular 911 calls; 6 (14%) felt it was *much better;* 5 (12%) felt it was *somewhat better;* 1 (2%) felt it was *somewhat* worse; and 1 (2%) felt it was *much worse than* typical cellular 911 calls. A chi-square analysis of responses revealed that there were significant differences in responses (x^2 = 65.62; df = 4; $p \leq .001$), with most respondents finding the location information *about the same* as location information typically provided by cellular 911 calls.

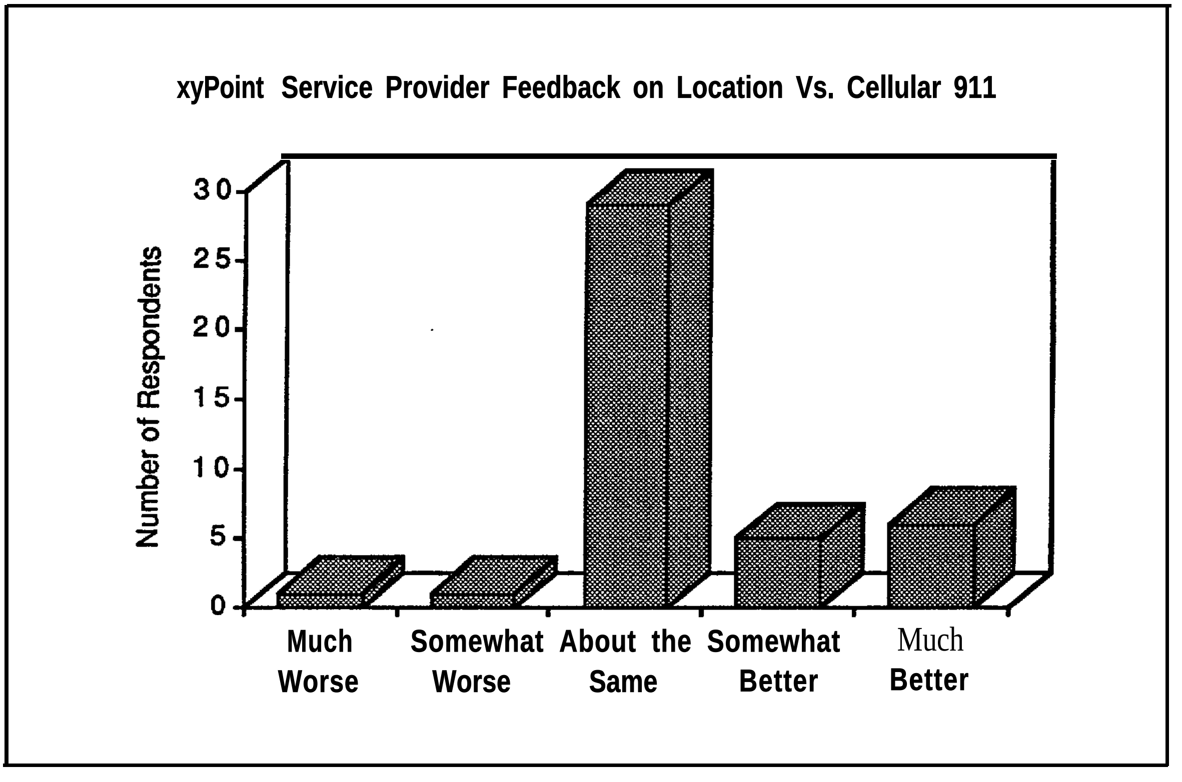


Figure 2.46 - Simulated Service Provider Attitudes Towards XYPOINT' s Location Information Compared to Information Received from Cellular 911 Calls

2.3.7 CONCLUSIONS: QUALITY OF INFORMATION

The results of the simulated service delivery test were somewhat disappointing since both the dispatchers and service providers found the quality of the incident and location information they received to be about the same as that which they currently receive from cellular 911 calls. This is particularly surprising in the case of location information since cellular 9 11 calls did not employ location technology such as GPS. However, this result was likely caused by factors other than the performance of the technology, specifically (1) the simulated nature of the incidents, and (2) the level of training and experience of the PuSHMe CSC operators. It should also be noted that aside from this comparison with 911 calls, the dispatchers and service providers tended to find the incident information useful; the dispatchers tended to find the location information reasonably accurate; and the service providers tended to find the location information reasonably to very accurate.

Survey comments of dispatchers and service providers were instructive. While some of their concerns could be related to performance (e.g. they would have liked to have had more descriptive location information, such as the comer of an intersection where the vehicle was located) most comments indicated that they expected the CSC operators to provide a greater level of details about the incident (e.g. suspect information, if incident was in progress, car description, etc.). However, the nature of the simulation precluded these details. The simulated service delivery scenarios could have included more "playing out" of the incidents, including details like suspect information and incident status, but more realistic role-playing was determined to be inappropriate. "Nor would it have been feasible," wrote a King County Police dispatcher, "too time consuming in an operating center, not to mention the confusion on the live radio frequency " The Task 2 Technical Memorandum (David Evans and Associates, August, 1996) provides additional reasons why detailed information was not provided:

> Long, detailed scenarios, conducted under low stress situations, could potentially cloud the performance evaluation of the system by providing skewed time data. Second, information about the suspect would be conveyed to the officers over radio which is heard by other non-participating agencies and could be acted on improperly. Third, the agencies in the focus groups and individual interviews indicated that in such an event (where a large amount of information would need to be exchanged) the caller should be patched directly to the service provider, a feature neither service could support at the time of testing. (p.63)

Another factor was the level of training and expertise of the CSC operators. While these operators did receive some system and protocol training from DEA and XYPOINT staff, they were not professionally trained to handle emergency calls and therefore were not intimately aware of the expectations of dispatchers and service providers. "King County Police spend months training their new employees to do this" commented a King County Police dispatcher. The Task 2 Technical Memorandum (David Evans and Associates, August, 19%) describes differences between industry training practices and operator training for the PuSHMe Operational Test.

2.4 SPECIFIC TEST 1: DROPPED CARRIER TEST

The first of our series of specific tests evaluated the ability of the PuSHMe system to handle dropped calls. Because calls to the Motorola response center could be disconnected by the user, unit malfunction, bad cellular connection, or phone line problems, Motorola built into the system an automatic re-dial feature that enables a dropped call to be automatically reconnected. This test evaluated the effectiveness and functionality of this feature.

2.4.1 OBJECTIVE OF TEST

The dropped carrier test evaluated the ability of the Motorola system to reconnect and handle a call after a connection had been terminated. It focused on whether the call connected with one button push by the in-vehicle user, whether the call automatically re-dialed the response center after being disconnected, and whether the call was then appropriately reconnected (i.e. recognized as a previously received and disconnected call).

2.4.2 TEST METHODOLOGY

The Dropped Carrier Test began with the user initiating a call. After the call was received at the CSC and acknowledged, the user terminated the call by disconnecting the unit's antenna. The user then reconnected the antenna. The call should then have automatically been re-established and operators at the CSC should have been able to identify them as previously received dropped calls. Specifically, if a call dropped "accidentally," the icon for the vehicle should have remained on the computer screen, but exhibited a different color. When reestablished, the call should have been classified under the same call identification number. Calls were considered validly connected when they were recognized by the system as a dropped call and reconnected under the same call identification number.

To supplement this additional test, trials from the User Croup Deployment (see Section 2.2.5.2.2) were analyzed to determine how often calls dropped during trials, and how often those dropped calls were validly reconnected. For each trial in the User Group Deployment, users indicated whether their call was disconnected and, if so, whether their call was then automatically reconnected.

2.43 DATA ANALYSIS

For the Dropped Carrier Test, calls were evaluated for reconnecting automatically and for reconnecting validly (i.e. under the same identification number), The trials were conducted by DEA and WSDOT staff. For each trial, DEA and WSDOT staff completed a log. The log indicated whether the call connected on the first button push, whether the call reconnected, and whether the call was assigned the same identification number. The logs of trials were then provided to the evaluation team.

For analysis of the User Group Deployment data, user report form data was linked to CSC workstation data and analyzed to determine how many calls were dropped and whether the dropped calls were re-established under the same call identification number.

2.4.4 RESULTS

2.4.4.1 Controlled Dropped Carrier Testing Results

A total of 150 trials were conducted. As Table 2.5 shows, 95% of calls connected with the first button push. Of those that connected with the first button push and were then disconnected, 93% of calls reconnected. Of calls that reconnected, 98% were reconnected validly (recognized as dropped calls and r-e-established under the same identification number). In other words, of 150 trials, 130 (87%) performed completely as expected.

Table 2.5 - Results of Dropped Carrier Test (Motorola System Only)

	Number	Percentage
Connecting on First Button Push	143/150	95%
Of Those Connecting, Those Reconnecting	133/143	93%
Of Those Reconnecting, Valid Reconnects (Calls Recognized as Previously Dropped Call)	130/133	98%

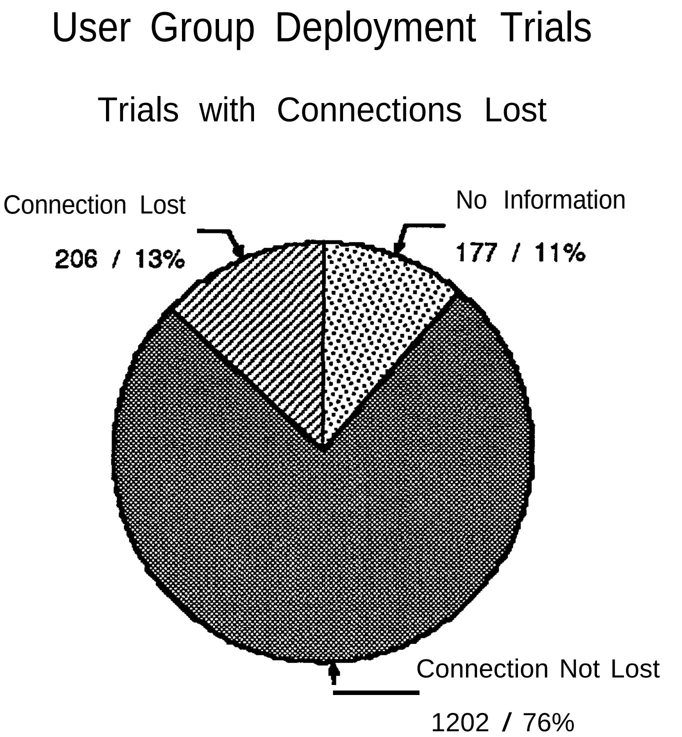


Figure 2.47 - User Group Deployment Trials with Lost Connections

2.4.4.2 Applicable Results From User Group Deployment

As shown in Figure 2.47, during the User Group Deployment 206 calls (approximately 13%) were disconnected during trials (as indicated on user report forms). Of these disconnected calls, users indicated that about 8 1% were automatically reconnected. However, from a user's perspective there was no distinction between simply being reconnected and being validly reconnected (i.e. assigned the same identification number).

To further understand this distinction between reconnection and ***valid*** reconnection, additional analysis was made relating (1) calls that lost contact during a trial to (2) trials that had multiple identification numbers. This analysis was based on the fact that users were instructed to conduct only one test a day. If a vehicle registered more than one call ID during a day, either (1) the user conducted more than one test during the day, or (2) a dropped call was incorrectly reestablished as a new call. With the exception of several partner vehicles which were often involved in more than one test during the day, it was fairly unlikely that a user would accidentally conduct a second test.

Therefore the registration of more than one call identification number per vehicle during a single day indicated a possibility that a dropped call had not properly been reestablished.

Results of this analysis are presented in Figures 2.48 and 2.49. Figure 2.48 presents the number of trials which lost contact grouped by those with a single call ID and those with more than one call ID. Figure 2.49 presents the same data as percentages.

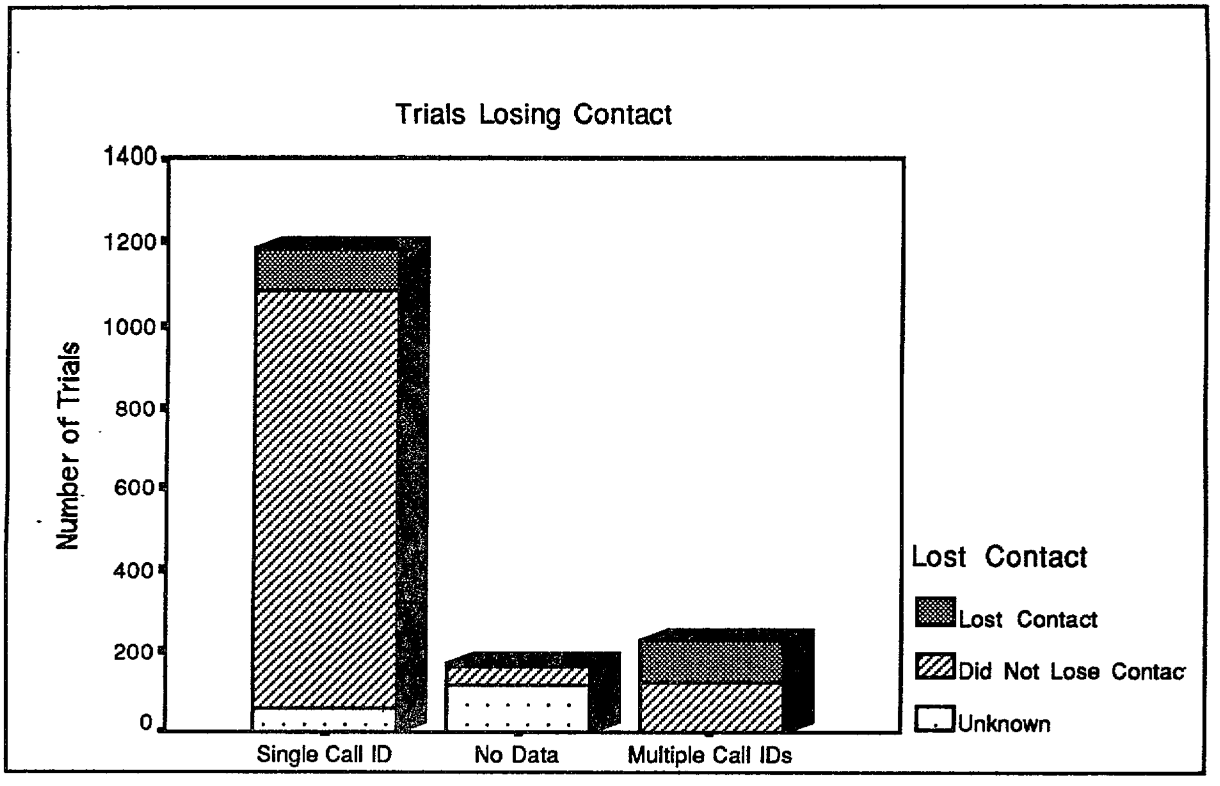


Figure 2.48 - Trials With Single or Multiple Call IDs - Number Losing Contact During Trial

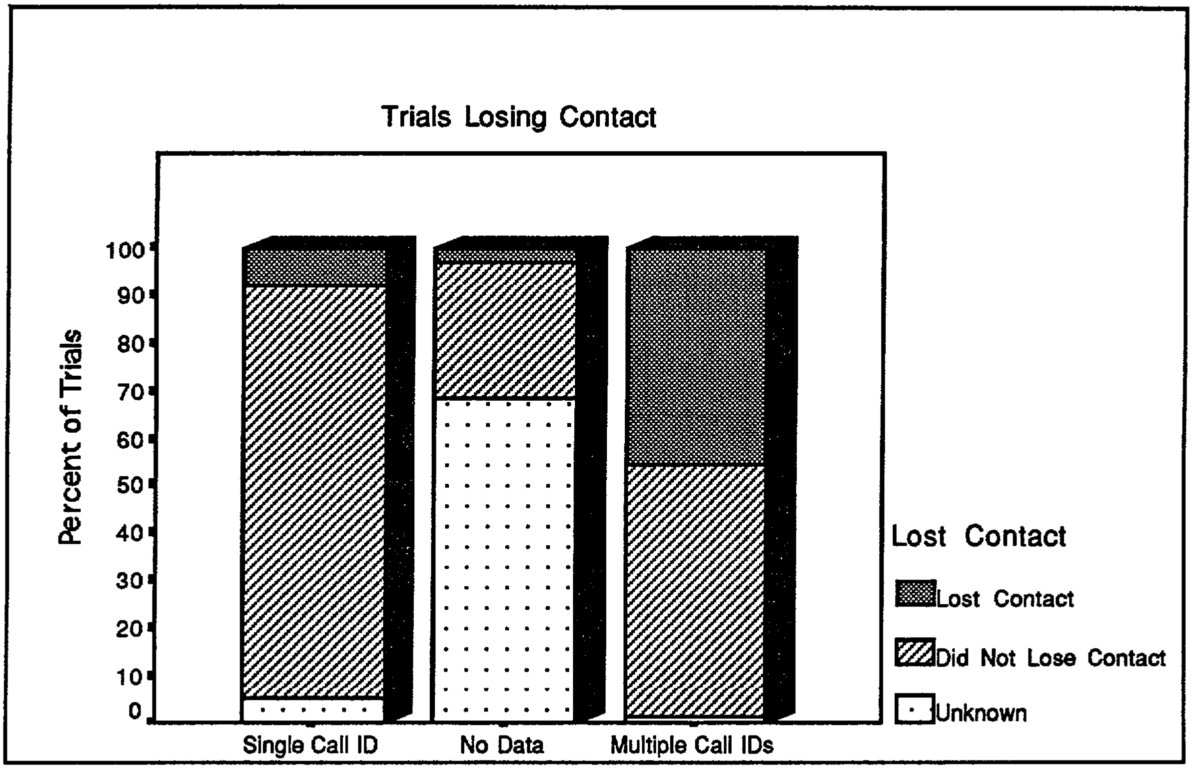


Figure 2.49 - Trials With Single or Multiple Call IDs - Percent Losing Contact During Trial

As shown in Figure 2.49, about half of the 228 trials with multiple call identification numbers had experienced a lost connection, indicating that about half were due to multiple tests from a single vehicle, rather than failure to validly reconnect. Along these lines it may be worth noting that about 60 (27%) of the trials with multiple call identification numbers were conducted by four partner vehicles involved in other system testing. The remaining 73% of trials with multiple call IDs were conducted by participants conducting only one test a day.

More importantly from a performance standpoint, of the 206 trials that users indicated had lost contact, 104 (50%) had multiple call ID numbers in the CSC data, indicating they had not validly reconnected. (This assumes that if a single trial included both a disconnected call *and* multiple IDs were associated with that trial, then the trial included a call which disconnected and then failed to validly reconnect.)

2.4.5 DISCUSSION AND CONCLUSIONS

The results of the Dropped Carrier Test demonstrate that the re-dial feature of the Motorola system is fairly successful at automatically reconnecting dropped calls in a controlled experiment. However, results from the User Group Deployment as evidenced by the rate of dropped calls exhibiting multiple call identification numbers (about 50%) also suggest that in ‘practice” dropped calls were not reliably reconnected within the context of CSC operations. In all likelihood, as in previous tests, operator error comes into play. During the user group deployment test, CSC operators received no instruction on the proper handling of an icon changing color (i.e. a call disconnecting) and an inappropriate action (e.g. prematurely deleting the icon) would result in a call receiving a second ID, rather than a valid reconnection. Given the extreme differences in results between the controlled Dropped Carrier Test and User Group Deployment analysis, this would seem to have occurred fairly often.

The proper handling of dropped calls would become particularly important during deployment since users’ calls would be placed in a queue when the response center is at capacity. The valid reconnections of dropped calls would ensure that a dropped call is re-established at its same rank, rather than at the end of the queue. Given the more than 1 in 10 trials that included calls where contact was lost, the importance of assuring that these “emergencies” don’t have to “start over” in seeking help seems extremely important.

Assuming that operator performance accounted for the wide disparity in the success of the re-dial feature between the Dropped Carrier and User Group Deployment tests, there seem to be two lessons to be learned: (1) the CSC interface design should be reevaluated to determine if it needs to be modified to enable better handling of dropped calls, and (2) attention must be paid to operator training to ensure that operators are familiar with the procedures required to properly handle dropped calls.

2.5 SPECIFIC TEST 2: TOPOGRAPHIC INTERFERENCE TEST

The second of our series of specific tests evaluated the performance of the GPS component of the PuSHMe system under various topographic conditions.

2.5.1 OBJECTIVE

Physical obstructions (such as tall buildings, overpasses, dense forests, etc.) can hinder a GPS receiver's ability to receive satellite signals. To determine how well the mayday systems performed in adverse or challenging locations, the evaluation team and the EWG conducted trials of the systems in locations characterized by overhead structures that could interfere with the reception of satellite signals.

2.52 TEST METHODOLOGY

The topographic interference test was conducted from January 22-25, 1996. The users initiated calls at prearranged locations that exemplified one of the following interference conditions: (1) in between buildings, (2) in a parking garage, (3) in a forest, and (4) no discernible barrier. Trials in parking garages were conducted on January 22, trials in between buildings were conducted on January 23, trials in wooded areas were conducted on January 24, and trials in open terrain were conducted on January 25.

Determining the precise accuracy of locations was not a specific feature of this test, but an effort was made to provide a general description of the quality of the data.

2.5.3 DATA ANALYSIS

The assignment of quality to the location data as listed below (Good, Close, Bad) was based on subjective ratings of the users in vehicles (DEA and WSDOT employees) in response to locations reported to them by CSC operators. These ratings may vary across users. There was also some inconsistency in how information was reported on the user logs. Some of the trial logs rated calls only as Good or Bad. Some of the trial logs did not include the assignment of a rating, but instead listed the address as reported to them by the CSC. When addresses, rather than ratings, were provided, the evaluators assigned a rating by comparing the location identified to the actual location of the user. Good was defined as within 1 block, close was defined as within 2 blocks, and bad was defined as more than 2 blocks or no location given.

In addition, the evaluation team reviewed the CSC data that was collected to determine whether GPS locations were being updated during trials. Even if vehicles are not moving, the GPS-determined locations should change slightly with each data record. When the location cannot be updated (e.g. because of poor satellite visibility) the GPS locations will remain exactly the same across records, since the "last known location" will be used in each subsequent record.

2.5.4 RESULTS

Tables 2.6 and 2.7 below present the results of the topographic interference tests. Trials conducted in each barrier type are discussed in more detail in section 2.5.5. In the tables below, the percent good, close, and bad arc taken from calls where connections were made.

Table 2.6 - Results of Topographic Interference Test (XYPOINT)

Barrier Type	Total Trials	Connecting	% Connecting	Good GPS	Close GPS	Bad GPS	Unknown	Locations Unable to Update
Garage	63	56	89%	0%	0%	100%	0%	100%
Between Buildings	38	31	82%	6%	23%	71%	0%	37%
Forest	25	23	92%	0%	83%	17%	0%	0%
Open	49	47	96%	40%	32%	19%	9%	8%

Table 2.7 - Results of Topographic Interference Test (Motorola)

Barrier Type	Total Trials	Connecting	% Connecting	Good GPS	Close GPS	Bad GPS	Unknown	Locations Unable to Update
Garage	43	39	91%	92%†	0%	8%	0%	100%
Between Buildings	45	38	84%	24%	26%	47%	0%	29%
Forest	25	23	92%	0%	96%	4%	0%	0%
Open	45	42	93%	55%	0%	45%	0%	44%*

t This 92% "good" result is not meaningful. See Section 2.5.5.1

*Note that for one unit tested on one day, locations were not updating for all trials. For other vehicle, all locations were updating.

2.5.5 RESULTS BY TYPE OF BARRIER

2.5.5.1 Trials Conducted in Parking Garages

For both the XYPOINT and Motorola systems, locations were unable to be updated for all trials conducted in parking garages. These results confirmed the obvious--GPS-based systems will experience difficulties in accurately determining locations in enclosed spaces like parking garages.

The previous table indicating that all XYPOINT trials were "bad" and most of the Motorola trials were "good" simply means that in the case of Motorola, the automatically reported "last known location” was very close to the actual location of the trial, while in the case of XYPOINT it was not. This was caused by different start-up procedures for the two systems and has nothing to do with the quality of performance.

2.5.5.2 Trials Conducted in “Urban Canyons”

For both XYPOINT and Motorola, systems experienced some trouble updating locations (37% and 29% failure to update respectively) in-between buildings (i.e., in “urban canyons”). The systems also experienced difficulties determining accurate locations (71% “bad” for XYPOINT and 47% “bad” for Motorola). These results confirm what was expected: GPS-based systems can experience difficulties in accurately determining locations in “urban canyons.”

2.5.5.3 Trials Conducted in Forests

For trials conducted in forests, both systems had no trouble updating locations. Most of the calls were ranked as “close” and none were ranked as “good.” The fact that most trials were “close” rather than “good” may relate to the on-line maps. The maps used on the terminals include most, but not all roads. Smaller roads are less likely than major roads to be included on the maps. For most trials placed in forests, then, the roads the vehicles were on were generally smaller roads that may not have been well-represented on the maps.

2.5.5.4 Trials Conducted in Open Terrain

For trials conducted in open terrain, the XYPOINT system experienced minor problems with updating location (8% failure) while Motorola appeared to have considerable difficulty (44% failure). However for the Motorola system, all trials conducted by one of the two units resulted in “bad” GPS locations, and a review of the CSC data indicated that for all trials of that unit, the GPS locations were not updating. For the other unit, all locations updated correctly.

2.5.6 DISCUSSION AND CONCLUSIONS

As was expected, both systems experienced difficulties in conditions with overhead obstructions. The most challenging overhead obstruction was parking garages. The second most challenging obstruction was in between buildings. Surprisingly, one of two Motorola units failed to work in open terrain. This is a reminder that the PuSHMe technology, while well understood, still needs to be “bullet proofed” before going to market. (Another reminder of this was the discovery, late in the project, that the differential correction component of the Motorola system had not be functioning properly. See 2.7)

2.6 SPECIFIC TEST 3: MOVING VEHICLE TEST

The third specific test evaluated the ability of the systems to track vehicles in motion.

2.6.1 OBJECTIVE

The objective of the moving vehicle test was to determine whether the PuSHMe systems were able to accurately track moving vehicles for a given period of time. The location accuracy of the systems is a function of at least two factors: (1) the ability of the GPS system to correctly locate moving vehicles, and (2) the accuracy of the on-screen maps.

2.6.2 TEST METHODOLOGY AND DATA ANALYSIS

Moving vehicle trials were conducted for both the Motorola and XYPOINT systems. The tests for the two systems were conducted using very different methods (because of the inherent differences in the technologies), so it would be misleading to compare the results for the two devices.

2.6.2.1 Test Methodology and Data Analysis for XYPOINT

Six XYPOINT units were deployed on February 14, 1996 to measure the ability of the XYPOINT system to track moving vehicles. Calls were made during rainy conditions in vehicles driving around a set route (the I-WI-405 loop).

The CSC operator tracked trial calls for 15 minutes, polling each unit approximately every 2 minutes. At the end of the trial, maps were generated showing the movement of the vehicle through the XYPOINT system. Thirty six trials were conducted resulting in 95 map printouts of data.

The printouts could be reviewed to evaluate whether the map displays a vehicle on or off the freeway, but it was not possible to determine how accurately the printouts locate a vehicle at a given milepost ***along*** a freeway at a given time. The fourth specific test, the location specific test (2.7), provides more detailed evaluation of the location accuracy of the systems.

2.6.2.2 Test Methodology and Data Analysis for Motorola

Two Motorola units were deployed over four days (February 5-8, 1996) to measure the ability of the Motorola system to track moving vehicles. Calls were made during rainy conditions in vehicles traveling along a pre-determined route (the 1-5/I-405 loop).

Users in the vehicles completed logs identifying the date, type of test, location, weather, and unit identification number. The users (DEA and WSDOT staff) also indicated the trial number, time of button push, whether a connection was made, time of connection, and their own subjective assessment of the general location accuracy of the data as quoted by the CSC operator.

Data was ranked as Excellent, Good, Fair, and Poor. Excellent indicated that the location was accurate for more than 80% of the trial, Good was accuracy for 60 to 80% of the trial, Fair was accuracy for 40 to 60% of the trial, and Poor was accuracy for less than 40% of the trial. The users in the vehicles were in constant communication with the CSC operators, so were primarily being told where along the freeway the CSC map located the user. Users also made note of instances when the computer at the CSC appeared to freeze or have other difficulties.

2.6.3 RESULTS

2.6.3.1 XYPOINT Moving Vehicle Test Results

Following is an assessment of the location accuracy (relative to the freeway) for moving vehicle trials for the XYPOINT system It represents the subjective perception of the evaluation team made by comparing the pre-determined route to the maps produced at the end of trials.

Table 2.8 Moving Vehicle Tests - XYPOINT

Location Accuracy	Number of Maps	Percent
Accurate within a half block	50	53%
Off by half a block	23	24%
Off by one block	17	18%
Off by two or more blocks	5	5%
Total	**95**	**100%**

***Represents location accuracy on or off the freeway, but does not include a determination of the accuracy of vehicles along a freeway at a given time**

2.6.3.2 Motorola Moving Vehicle Test Results

Following are the subjective ranking of trials as made by the DEA and WSDOT participants in the moving vehicle tests:

Table 2.9 - Moving Vehicle Tests - Motorola

Location Accuracy	Number of Trials	Percent
Computer Problem (Froze, Had to Reboot, No Icon, etc.)	13	13%
No Connect	6	6%
Excellent	33	32%
Good	44	43%
Fair	3	3%
Poor	2	2%
No Rating	1	1%
Total	102	100%

Of trials that connected and did not experience computer workstation problems, 33 of 83 (40%) were rated as excellent location and 44 of 83 (53%) were rated as good location.

2.6.4 DISCUSSION AND CONCLUSIONS

Both PuSHMe systems were able to track moving vehicles reasonably well, although each had its drawbacks. In the case of the XYPOINT system the issue was process since it was highly labor intensive for the CSC operator to continuously poll the vehicle in order to track it. In the case of the Motorola system, as in the Topographic Interference Test, the issue was "bullet proofing" as nearly one in five trials never connected.

2.7 SPECIFIC TEST 4: LOCATION SPECIFIC TEST

The fourth specific test evaluated the positioning accuracy of the two PuSHMe systems

2.7.1 OBJECTIVE AND BACKGROUND

Both the Motorola and XYPOINT systems rely on the United States Department of Defense (USDOD) Standard Positioning System (SPS), which is available to all users worldwide. While it goes beyond the scope of this evaluation to provide a complete description of the Global Positioning System (GPS), the results of the location specific test can best be understood within the context of the typical location accuracy that can be expected with the use of GPS.

The SPS provides predicable accuracy of 100 meters 95% of the time in the horizontal plane. This accuracy is established by both the USDOD and the USDOT, and is based on U.S. security interests (Kaplan, 1996, p. 5). In actual use, however, the typical accuracy may be better than 100 meters. At MIT/Lincoln Laboratory, an FAA-sponsored program has examined the level of performance achievable using undifferentially corrected GPS with selective availability on and the Russian-owned GLONASS system (Misra, 1993). This program has found that at MIT/Lincoln Laboratory the typical error has been in the range of 25 to 40 meters. (For more information about the Lincoln Laboratory project, see Misra, 1993 or the MIT/Lincoln Laboratory GLONASS Group web pages--http://satnav.atc.lll.mid.edu; and http://satnav.atc.lll.mid.edu/gps/images/gps-scat.gif).

To improve accuracy, both the Motorola and XYPOINT system used differentially corrected GPS (DGPS). The Motorola data was differentially corrected at the Washington State Traffic Systems Management Center (TSMC), which was equipped with a DGPS base station by PuSHMe. The XYPOINT data was differentially corrected at Trimble's base station in Lynnwood, Washington. The use of DGPS enhances GPS accuracy to produce typical position errors of less than 10 meters (Kaplan, 1996, p.322).

2.7.2 METHODOLOGY

For the PuSHMe location specific test, users were instructed to drive to specific location markers (with known location coordinates), park the car, and initiate a trial. When users were unable to park their car directly over the monument, they noted approximately how far away they were from the monument. With the button push, data was sent to the CSCs indicating the location of the vehicle. For the Motorola system, the location was updated every 5 seconds even though the vehicle did not move. For the XYPOINT system, updates occurred with each button push. Therefore, for each trial multiple locations were sometimes identified (particularly with the Motorola system). For this analysis, the locations generated for a single trial were averaged. For trials in which users indicated that they were off the monument, the evaluators gave the PuSHMe systems the benefit of the doubt when the location data was analyzed. The evaluators compared the two sets of coordinates (the known coordinates to those determined by the system), and determined how accurate the CSCs were in precisely locating vehicles.

2.7.3 DATA ANALYSIS

Comparing location coordinates to determine location accuracy is a somewhat complicated process and involves converting data as many as three times: (1) the two measurements (CSC data and "known location" data) must be in the same "physical" format (decimal versus geodetic), (2) the CSC data and "known location" data need to be based on the same (or comparable) geodetic coordinate systems, and (3) to determine the difference between two locations in meters, the geodetic coordinates can be converted to the state plane coordinate system (See Appendix E for an explanation of how data was converted

2.7.4 RESULTS

Figures 2.50 and 2.51 present results of the Location Specific Test for Motorola and XYPOINT, respectively. Each figure presents a histogram of how many meters individual trials were off of the monument, and on the right side of the figure, the cumulative percent of trails. Figure 2.52 is a scatter diagram indicating the system derived locations relative to the known locations of the monument.

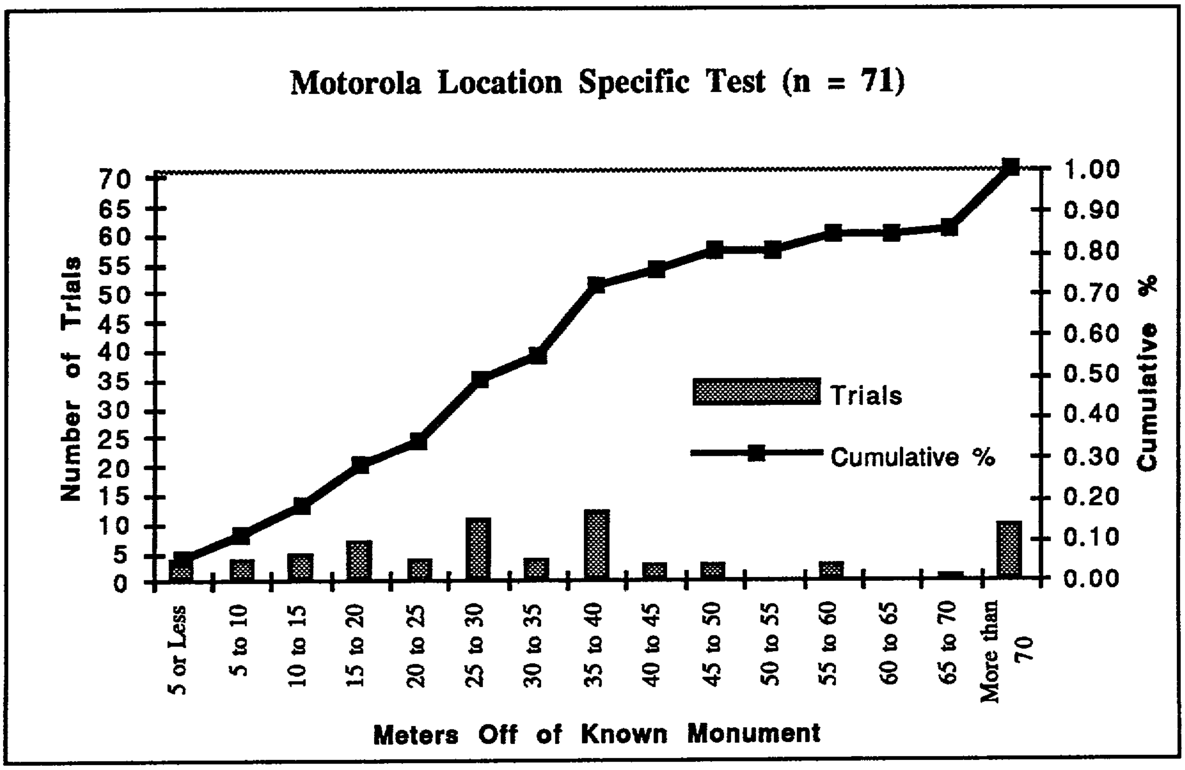


Figure 2.50 - Location Specific Test Results - Motorola Histogram

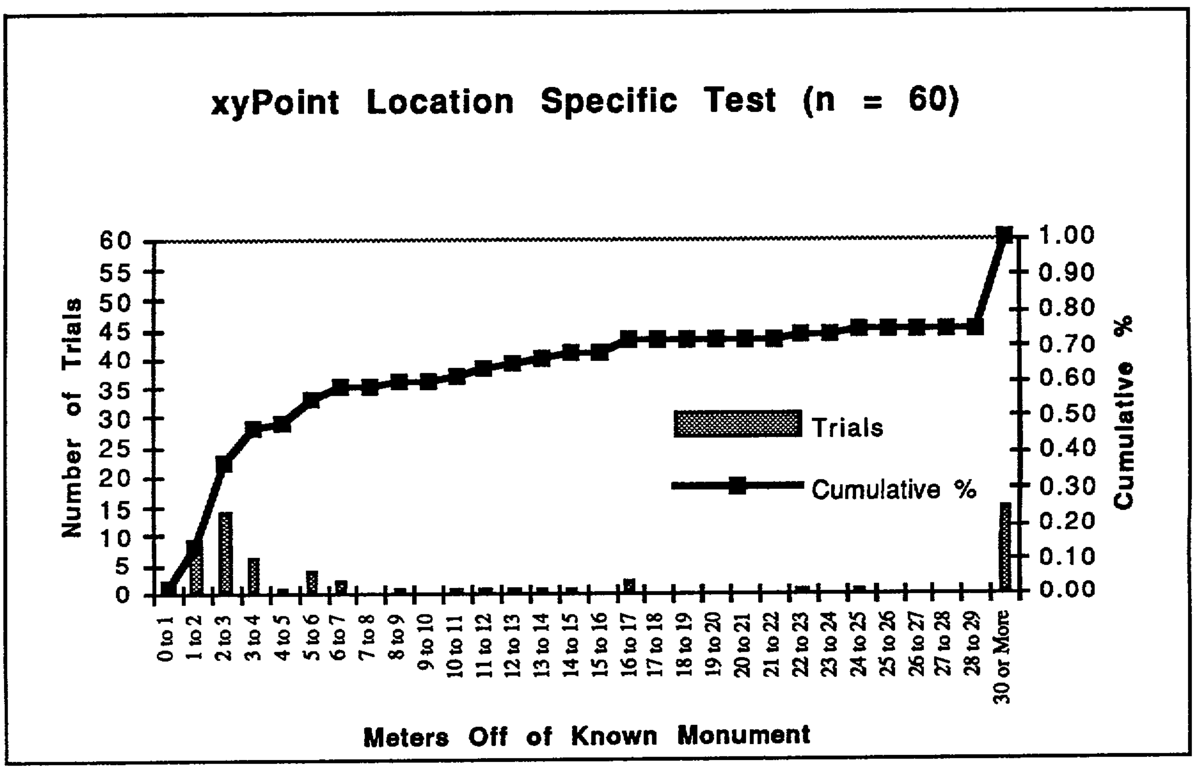


Figure 2.51 - Location Specific Test Results - XYPOINT Histogram

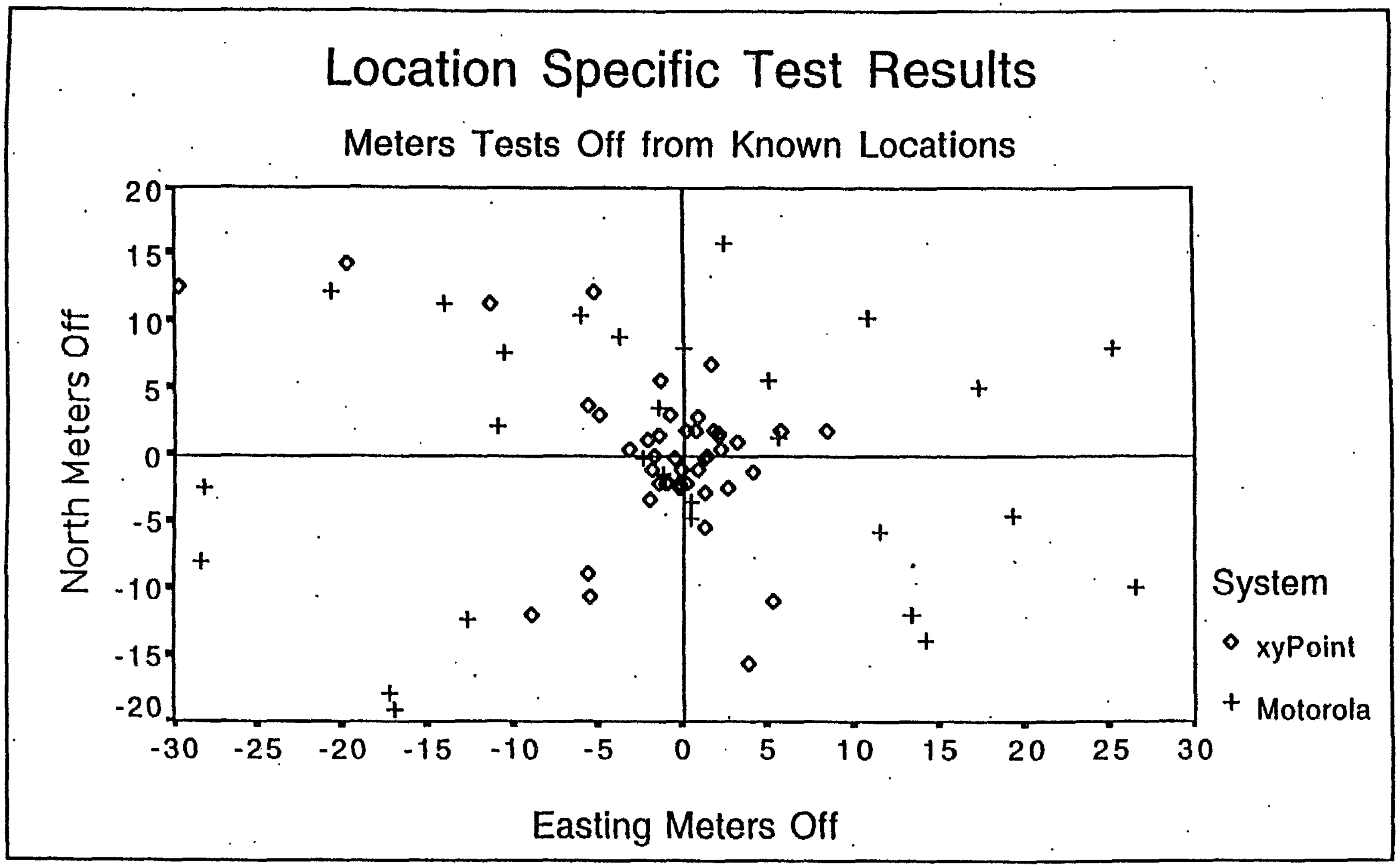


Figure 2.52 - Location Specific Test Results - Scatter Diagram of 30 Meters Away from Monument

As shown in these figures, most of the Motorola locations were within 30 meters of the monument, and most of the XYPOINT locations were within 6 meters of the monument.

2.7.5 DISCUSSION AND CONCLUSIONS

The Motorola trials produced location accuracy typical of undifferentially corrected GPS. While the base station at the Washington State Transportation Systems Management Center (TSMC) was supposed to be differentially correcting Motorola locations, these results (plus subsequent discussion with the partners) indicate that the TSMC base station was probably not functioning for all or part of the PuSHMe test. The mean distance off for Motorola data was about 37 meters, and the median distance off was about 31 meters. These results are comparable to undifferentially corrected GPS results being obtained by the MIT/Lincoln Laboratory (see Section 2.7.1), which found typical errors in the 25 to 40 meter range.

The XYPOINT trials produced location accuracy typical (or slightly better than typical) of DGPS. However XYPOINT produced a relatively large number of "outlier" locations with several trials more than 80 meters off of the known location. The cause of these outliers is not clear. Locations off by 80 or more meters could be explained by undifferentially corrected data, however for all of the trials off by more than 80 meters, the CSC data indicated that the locations were being differentially corrected Other possible explanations for extreme outliers include: incorrect logging of trial by user or operator, topographical interference, and hardware or software failure.

The adequacy of various location accuracies for emergency response services is a highly situational question. If the goal is to know in which lane a vehicle is sitting, a 6 meter error may be too large; if the goal is to find a stranded vehicle on a rural highway, an 80 meter error may be tolerable. In general, both PuSHMe systems produced relatively reliable, relatively accurate locations that in most cases would be helpful in support of the delivery of emergency services.

2.8 SPECIFIC TEST 5: REMOTE CSC OPERATOR TEST

The fifth and final specific test evaluated the ability of a remote CSC to determine the location of a vehicle in the Seattle area.

2.8.1 OBJECTIVE

Different emergency response systems will encounter peak loads at different times or in different situations. For example, a large-scale traffic accident or a natural disaster could overload a local processing center. In these situations, service centers in other areas of the country could be used to process calls when local service center operators are overloaded or unable to answer calls. The Remote CSC Operator Test (referred to in the Final Detailed Evaluation Plan as the Seattle to Phoenix Test) evaluated one component of remote emergency management--the ability of a remote operator to locate the distressed vehicle.

2.8.2 TEST METHODOLOGY

For the Remote CSC Operator Tests, DEA staff conducted in-vehicle trials in Seattle and the calls were received by CSC operators (Motorola staff) in Phoenix, Arizona. One unit was deployed to locations within the Seattle Area (Ring and Snohomish Counties) over a two day period (May 22-23, 1996) The DEA tester drove to a random location; parked the car (except for one trial on the I-90 bridge); initiated a trial call; and filled out a log sheet identifying date, trial number, location, time initiated, time contacted, location verified, and button pushed. At the Phoenix location, the operator wrote down the location identified.

2.8.3 DATA ANALYSIS

For this test, the user-identified location and operator-identified location were compared to determine how successful the remote operators were in locating vehicles. Four ratings of success were identified: (1) exact location identified; (2) very close (location was within half a block, the street name was slightly inaccurate, or the operator identified a stationary vehicle as moving); (3) fairly close (off by a block or more); and (4) operator could not/did not identify one cross street.

2.8.4 RESULTS

The trial logs for the Remote CSC Operator test are included in Appendix F. As shown in Table 2.10, for 60% of trials, the remote CSC operator was able to identify the *exact location* of the vehicle. For 22% of trials the remote CSC operator was very close, either being within a half block, naming one of the cross streets slightly incorrectly, or showing a stopped vehicle as moving. Therefore, for 82% of the trials the remote CSC was able to identify the location of the vehicle within half a block. For one trial (2%), the operator-identified location was off by more than a block. In 16% of trials, the operator either did not or could not correctly identify the name of one of the cross streets (although for the most part they seemed to be able to identify the names of other streets or landmarks nearby).

Table 2.10 - Results of Remote CSC Operator Test

Rating	Percent of Trials	Number of Trials
Exact Location	60%	30
Very Close	22%	11
Fairly Close (Off by a Block or More)	2%	1
Did/Could Not Identify Street	16%	8

2.8.5 DISCUSSION AND CONCLUSIONS

The Remote Operator test was extremely successful, in fact the Motorola operator in Phoenix was generally more accurate in his location of PuSHMe vehicles than were CSC operators in Seattle for the Motorola part of the Location Specific test. For 82% of trials the Phoenix operator was able to locate the users quite accurately. For 16% of trials, the remote operator was unable to name one of the cross streets, but for the most part was able to identify other nearby cross streets or landmarks (e.g. railroad tracks, other streets, census tracts, beach, etc.).

The inability to name one of the cross streets may have been a mapping problem. The problem of incomplete or out-of-date maps is not unique, though perhaps more common, to the remote operator scenario. Clearly a part of the success of GPS-based emergency response systems depends upon the existence of very detailed, up-to-date maps. The availability of such maps on the WWW would allow for uniformity and cross referencing among geographically disperse CSCs.

The success of remote location does not automatically translate into successful remote emergency management. Numerous other issues such as extra-jurisdictional processing would need to be addressed before remote emergency management was deemed practical.

Chapter 2 has presented a detailed analysis of numerous performance aspects of the PuSHMe emergency response systems. Overall the systems performed extremely well, though in many cases were not ready for "prime time." This is entirely appropriate since these systems were pre-market versions and no fundamental technological barriers were revealed that would prevent their being taken to market readiness.

System performance is a crucial issue, but it is not the only one. Chapter 3 turns to the issue of user acceptance and ease of use, or "usability."

CHAPTER 3. USABILITY ANALYSIS

3.1 OVERVIEW OF USABILITY ANALYSIS

The goal of the usability analysis was to evaluate the user's acceptance and ease of use of the PuSHMe mayday systems (at its current stage of development) by measuring the users' perceptions of the systems. The users involved in this evaluation were volunteers participating in the user group deployment (see Section 2.2.1). For several months, volunteers conducted daily tests of the system from their vehicles. During these tests, no actual service providers were dispatched to the volunteers.

3.2 METHODS

Two different questionnaires were administered: the first to volunteers using the Motorola system and the second to volunteers using the XYPOINT system (see Appendix G). The questionnaires explored four areas of user acceptance: (1) ease of use, (2) safety and security, (3) reliability and consistency, and (4) additional user perceptions of the testing or systems. A fifth section on the questionnaire requested information on drivers' driving habits and a limited amount of demographic data.

Before questionnaires were finalized, each draft questionnaire was pre-tested on five PuSHMe project team members from the partner organizations (DEA and WSDOT) The questionnaires were then revised based on feedback provided by the pre-testers. Once questionnaires were finalized, each volunteer was sent a questionnaire, a letter explaining the purpose of the questionnaire, and a postage paid return envelope to mail the completed questionnaire back to the evaluation team. Questionnaires were mailed to all volunteers near the end of the user group deployment. 36 questionnaires were sent to volunteers using the Motorola system and 83 questionnaires were sent to volunteers using the XYPOINT system For volunteers using the Motorola system, 23 completed questionnaires were returned, representing a response rate of 64%. For volunteers using the XYPOINT system, 54 completed questionnaires were returned, representing a response rate of 65%.

3.3 DEMOGRAPHIC CHARACTERISTICS OF RESPONDENTS

Ideally, the volunteers participating in the PuSHMe project would have been randomly selected from the population at large. However, as described in Section 1.4.2, because involvement in this project required an extensive time commitment, the participants could not be randomly selected. Instead, DEA recruited participants by identifying organizations that might be interested in in-vehicle mayday system technology and could support the test. DEA contacted these organizations and recruited volunteers who indicated that they (1) had a dedicated vehicle that could be used in the test; (2) were located in King and/or Snohomish counties within the cellular network; (3) were frequent travelers; and (4) had schedules that would provide availability during study hours.

Demographic data on volunteers were collected by both DEA and the evaluation team. DEA collected information on volunteers' age, education, income, average annual vehicle miles traveled, and occupation. The evaluation team collected information on volunteers' cellular phone ownership, use of pagers, and use of a home security system. The next two sections present the demographic characteristics of Motorola and XYPOINT volunteers who completed the usability surveys.

3.3.1 MOTOROLA RESPONDENTS

Of the Motorola volunteers completing the usability surveys, eighteen (81.8%) were male and four (18.2%) were female. The age of volunteers ranged from 28 to 64, the average age was 43 (SD = 8.7). The annual vehicle miles traveled (VMT) ranged from 5,000 to 40,000 miles; the average annual VMT was 16,000 (SD = 8,100). Twelve respondents (52.2%) indicated that they own a cellular phone; seven respondents (30.4%) indicated that they carry a pager; and six respondents (27.3%) indicated that their home is protected by a home security system.

Figures 3.1 through 3.3 present a breakdown of the education, income, and type of occupation of respondents. As shown in these figures, the majority of respondents are University or Community College educated; have incomes ranging from $30,000 to $75,000; and are in technical professions (e.g., technicians, engineers).

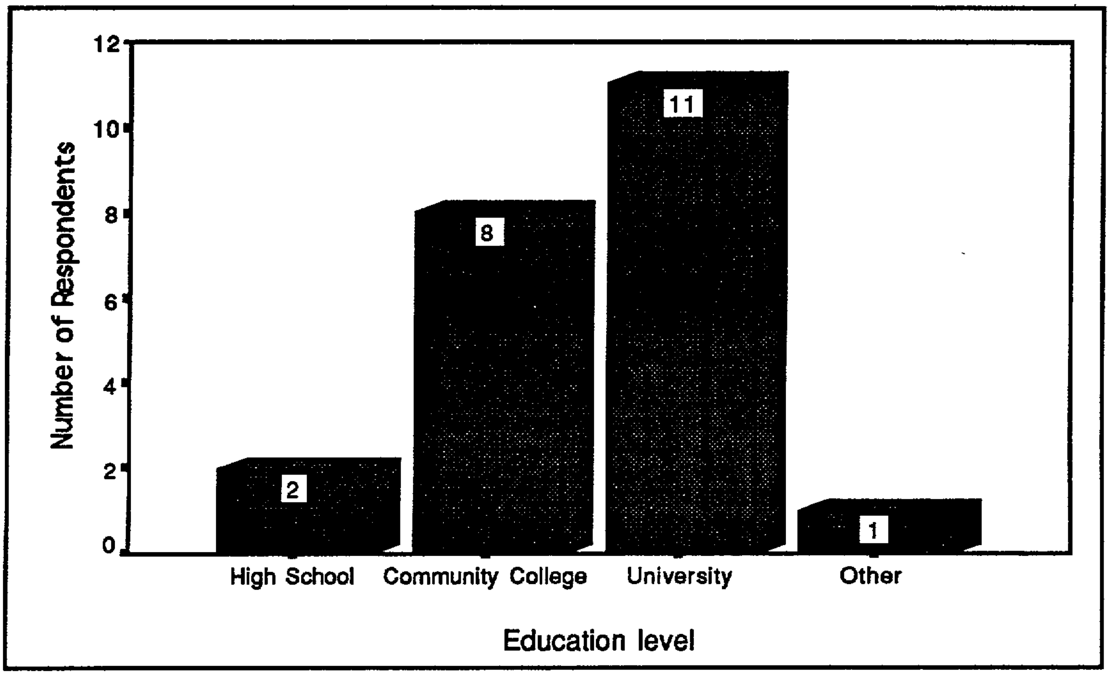


Figure 3.1 - Motorola Usability Respondents By Education
Numbers of respondents per category are displayed in bars

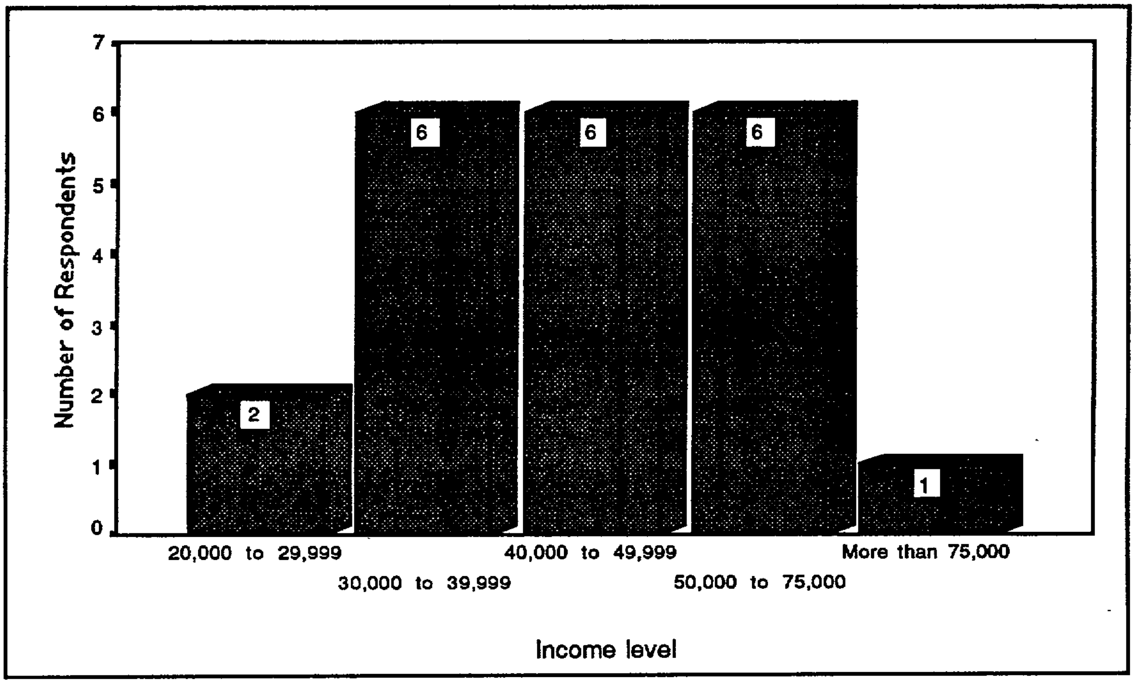


Figure 3.2 - Motorola Usability Respondents By Income
Numbers of respondents per category are displayed in bars

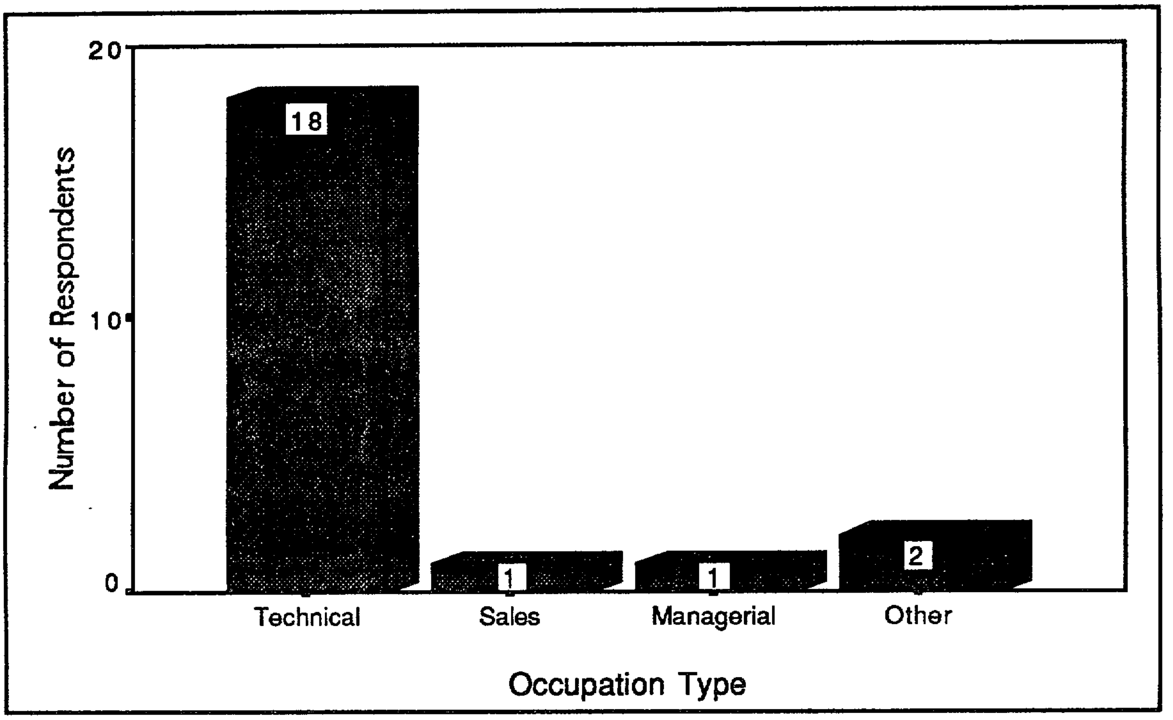


Figure 3.3 - Motorola Usability Respondents By Occupation
Numbers of respondents per category are displayed in bars

3.3.2 XYPOINT RESPONDENTS

Of the XYPOINT volunteers completing usability surveys, forty-two (82.4%) were male and nine (17.6%) were female. The age of volunteers ranged from 25 to 67; the average age was 45 (SD = 8.7). The annual vehicle miles traveled (VMT) ranged from 10,000 to 50,000 miles; the average annual VMT was 18,000 (SD = 7,800). Thirty-one respondents (58.5%) indicated that they own a cellular phone; nineteen respondents (35.8%) indicated that they carry a pager; and sixteen respondents (34.8%) indicated that their home is protected by a home security system.

Figures 3.4 through 3.6 present a breakdown of the education, income, and type of occupation of respondents. As shown in these figures, the majority of respondents are University or Community College educated; have incomes ranging from $30,000 to $75,000; and are in managerial or technical (e.g., technicians, engineers) professions.

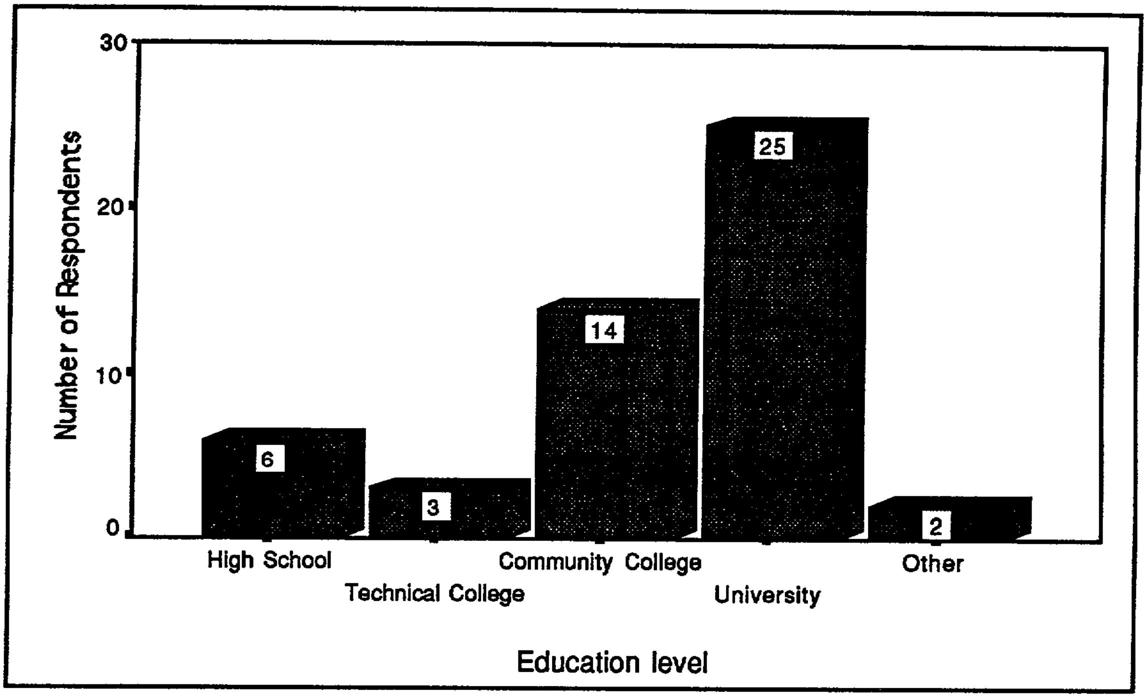


Figure 3.4 - XYPOINT Usability Respondents By Education
Numbers of respondents per category are displayed in bars

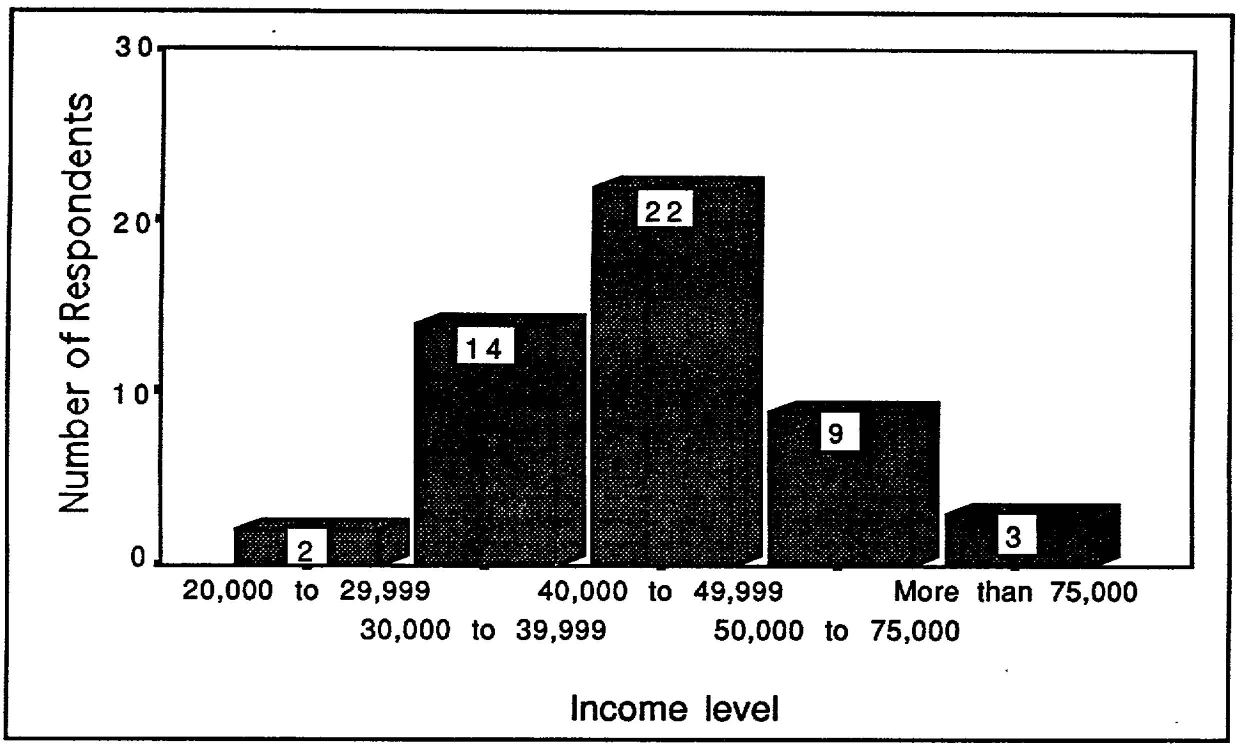


Figure 3.5 - XYPOINT Usability Respondents By Income
Numbers of respondents per category are displayed in bars

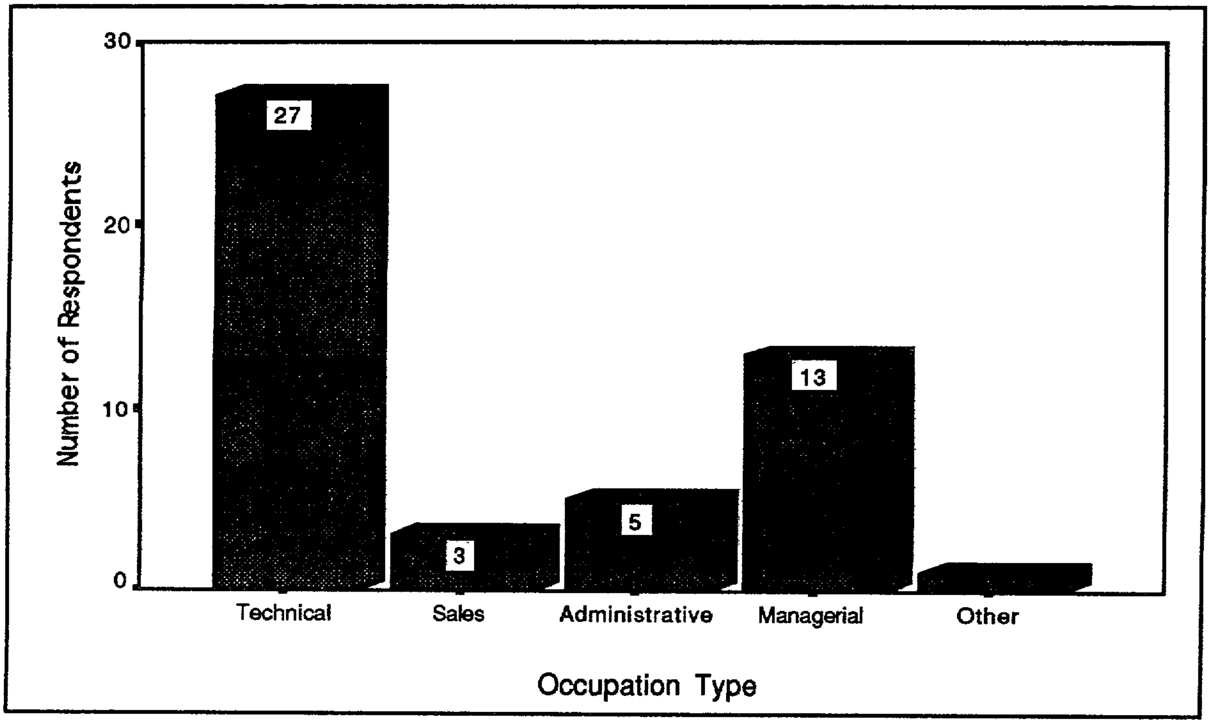


Figure 3.6 - XYPOINT Usability Respondents By Type Of Occupation
Numbers of respondents per category are displayed in bars

3.4 DATA ANALYSIS

Chi-square analyses were used to test for significant differences (at an alpha-level of .05) among responses on nominal scales in the survey questions. This statistical test ensured that differences in the responses were sufficiently great (when reported *p* values are less than or equal to .05) so as not to represent random variation. In addition, correlation analyses, with an alpha-level of .05, were used to determine if responses to individual questions significantly correlated with responses to other questions in the survey. ANOVAs with an alpha-level of .05 or less were used to determine if there were significant differences in responses to questions on interval scales. When question responses were on ordinal scales, Mann-Whitney tests were used (as opposed to t-tests which would have been used with interval scales) to assess group differences when there were only two groups. Similarly, Kruskal-Wallis tests were used to assess group differences (as opposed to ANOVAs) when there were more than two groups responding to questions based on ordinals. Some respondents did not answer all questions on the survey. For this reason, the total number of responses to some questions does not equal the sample totals (23 for Motorola and 54 for XYPOINT).

The results are presented by topic rather than by order of the questions on the survey. The tables summarizing the results include the original survey question number for reference, and copies of the surveys are included in Appendix G.

3.5 RESULTS

The survey results are separated by the system used. Section 3.5.1 presents results of the surveys administered to volunteers using the Motorola system and Section 3.5.2 presents results of the surveys administered to volunteers using the XYPOINT system.

3.5.1 MOTOROLA RESULTS

3.5.1.1 Motorola User Perceptions of Ease of Use: General Ease of Use

The survey included several questions that focused on respondents' perception of general ease of use of the device or system. Answers to these questions were generally very positive. When asked if they found the device easy to reach and easy to handle (questions 2a and 2b, Section l), all 23 respondents answered yes ($c^2 = 23.00$, $p < .01$; $c^2 = 23.00$, $p < .01$, respectively). Responses to additional questions relating to ease of use are presented in Table 3.1. As shown in Table 3.1, most respondents either *agreed* or strongly *agreed* that *they* found the device easy to use (100%), found the written instructions for using the device easy to understand (90%), knew what to expect when they operated the device (95.7%), found the auto re-dial feature useful (100%), and found the operator's voice at the response center easy to hear (78.2%) and easy to understand *(86.9%)*. Chi square analyses showed significant differences in responses for these results, with responses favoring the positive end of the scale. Further, respondents felt confident in selecting which button to push, *with 73.9%* indicating *that they strongly agreed,* and 26.1% indicating that they *agreed* ($c^2 = 33.52$, $p < .05$). The results of the user group deployment (see Section 2.2) indicate that, for most scenarios, there was consistency in the button users pushed. For 15 of 18 scenarios tested in the user group deployment, there was 90% or greater consistency in the button pushed by users (see Appendix H for detailed data).

Table 3.1- Motorola Perceptions of General Ease of User of the System

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	c^2	*p*
Found device easy to use (Sec. 1, 3g).	0.0% (0)	0.0% (0)	47.8% (11)	52.2% (12)	23.09	.0000
Found the written instructions for using the device easy to understand (Sec. 1, 3q).	10.0% (2)	0.0% (0)	70.0% (14)	20.0% (4)	23.20	.0000
Knew what to expect when operated device (Sec. 1, 3d).	4.3% (1)	0.0% (0)	56.5% (13)	39.1% (9)	20.65	.0001
Found the auto-redial feature useful (Sec. 1, 3f).	0.0% (0)	0.0% (0)	59.1% (13)	40.9% (9)	23.45	.0000
Found the operator's voice at the response center easy to hear (Sec. 1, 3l).	4.3% (1)	17.4% (4)	73.9% (17)	4.3% (1)	30.39	.0000
Found the operator's voice at the response center easy to understand (Sec. 1,3m).	0.0% (0)	13.0% (3)	65.2% (15)	21.7% (5)	22.04	.0001
Felt confident in selecting which button to push (Sec. 1, 3e).	0.0% (0)	0.0% (0)	26.1% (6)	73.9% (17)	33.52	.0000

Note: Number in parentheses under percentage shows cell size.

Many responses to these questions significantly correlated with responses to other survey questions. For example, the level of agreement on responses to the device being easy to use positively correlated with responses to finding the operator's voice easy to understand, and finding the auto-redial feature useful. The level of agreement on responses to participants knowing what to expect when they operated the device positively correlated with responses to feeling confident in selecting which button to push for each scenario they were given. This in turn positively correlated with responses to finding the device easy to use, and believing this system would be likely to help authorities deliver assistance when they are in situations requiring police, medical, or roadside assistance. The complete set of survey questions, with responses that significantly correlated with responses to the general ease of use questions, and the relevant statistics, appear in Appendix J.

Overall, the responses to questions focusing on user perceptions of general ease of use of the system indicate that in nearly every respect there was widespread satisfaction with the ease of use of the Motorola system.

3.5.1.2 Motorola User Perceptions of Ease of Use: Microphone Versus Handset

Because users could communicate with the operator using either a microphone or a handset (see Section 1.3.2 for a more complete description of the Motorola system), the usability questionnaire included several questions regarding ease of use and preferences for the microphone or handset.

As shown in Table 3.2, when using only the microphone, 63.2% of the respondents *almost* always heard the operator and 3 1.6% *frequently* heard the operator. These respondents, however, reported that the operator could not hear them as well: only 35.3% of respondents indicated that the operator seemed to *almost always* hear them when using only the microphone, although 41.2% indicated that the operator *frequently* heard them, and 23.5% indicated that the operator *occasionally* heard them.

When using only the handset, 82.4% of respondents *almost* always heard the operator and 11.8% *frequently* heard the operator. In contrast to those who used the microphone, the handset users reported that the operator seemed to be able to hear them well: 82.4% of respondents indicated that the operator seemed to *almost always* hear them when using only the handset, 11.8% indicated that the operator *frequently* heard them, and only 5.9% indicated that the operator *rarely* heard them.

Table 3.2 - Motorola User Perceptions of Ease of Use of Handset Versus Microphone

Question	Rarely	Occasionally	Frequently	Almost Always	c^2	*p*
When used only microphone, could hear the operator (Sec. 1, lb).	0.0% (0)	5.3% (1)	31.6% (6)	63.2% (12)	**19.11**	.0003
When used only microphone, was heard by operator (Sec. 1, lc).	0.0% (0)	23.5% (4)	41.2% (7)	35.3% (6)	6.76	.08
When used only handset, could hear the operator (Sec. 1, Id).	0.0% (0)	5.9% (1)	11.8% (2)	82.4% (14)	**43.47**	.0000
When used only handset, was heard by operator (Sec. 1, le).	5.9% (1)	0.0% (0)	11.8% (2)	82.4% (14)	**30.29**	.0000

Note: Number in parentheses under percentage shows cell size.

When respondents used only the microphone, the chi square analysis for the respondents hearing the operator was significant (with the *almost always* response being selected by a majority of the respondents), yet the chi square analysis missed being significant ($p = .08$) for the operator hearing the respondent-an effect probably caused by the relatively even response *to the almost always* and *frequently* responses. When respondents used only the handset, the chi square analyses were significant for both the respondents hearing the operator and the operator hearing the respondent, *with the almost always* response being selected most often. The responses, then, suggest that the handset enabled respondents to better hear and be heard by the operator.

Although using the handset appears to lead to clearer reception, other features of the system apparently contribute to the communication mode (handset versus microphone) that users find most effective. As shown in Table 3.3, when respondents were asked if they found the handset more effective than the microphone, *43.8% agreed* and 18.8% *strongly agreed,* a total of 62.6%. When asked if they found the microphone more effective than the handset, 29.4% *agreed* and 35.3% *strongly agreed,* a total of 64.7%. Oddly, 62-65% of respondents *agreed or strongly agreed* on separate questions to finding the handset *and* the microphone more effective. However, when asked which they preferred to use, more respondents preferred the microphone: 70% *agreed* or *strongly agreed* that they preferred to use the microphone, while only 45.5% of respondents *agreed* or *strongly agreed* that they preferred to use the handset. The chi square analyses were not significant for any of the questions presented in Table Z-an effect caused by the relatively even distribution of responses.

Table 3.3 - Motorola User Preferences for Handset Versus Microphone

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	c^2	*p*
Preferred to use the handset (Sec. 1, 3h).	22.7% (5)	31.8% (7)	27.3% (6)	18.2% (4)	.91	.8232
Preferred to use the microphone (Sec. 1, 3i).	10.0% (2)	20.0% (4)	35.0% (7)	35.0% (7)	3.60	.3080
Found the handset more effective than the microphone (Sec. 1, 3j).	6.3% (1)	31.3% (5)	43.8% (7)	18.8% (3)	5.00	.1718
Found the microphone more effective than the handset (Sec. 1, 3k).	11.8% (2)	23.5% (4)	29.4% (5)	35.3% (6)	2.06	.5603

Note: Number in parentheses under percentage shows cell size.

Users of the microphone and handset could clearly hear the operator, although users of the handset were more likely to also be heard clearly by the operator. However, the users' preference and views of effectiveness between the microphone and handset did not differ significantly. Their views about the microphone and handset did, however, significantly correlate with the responses to other questions. For example, the level of agreement on responses to the operator being able to hear participants clearly when using only the microphone positively correlated with responses to finding the microphone more effective than the handset, as well as preferring to use the microphone, and inversely correlated with responses to preferring to use the handset, as well as finding the handset more effective than the microphone. The complete set of survey questions with responses that significantly correlated with responses to questions regarding the ease of use of the microphone and handset, and the relevant statistics, appear in Appendix J.

Overall, most respondents reported that they could hear the operator and vice versa, although the handset appeared to lead to clearer reception (particularly with the operator being able to hear the user). Most respondents, however, still preferred to use the microphone, which enabled hands-fiee communication.

351.3 Motorola User Perceptions of Security/Safety

The questionnaire also included several questions regarding user perceptions of safety or security they believed this system could offer. Answers to these questions were very positive. As shown in Table 3.4, a large majority of respondents either *agreed or strongly agreed* that they would feel more secure in their vehicle if this system were permanently available to them (95.7%) and to other members of their family (95.7%), and that this system would be likely to help authorities deliver assistance when they are in situations requiring police, medical, or roadside assistance (95.6%). Chi square analyses showed significant differences in responses to these questions, with responses favoring the positive end of the scale.

Table 3.4 - Motorola User Perceptions of Safety and Security Benefits of System

Question	Strongly Disagree	Disagree	Agree	**Strongly Agree**	**c^2**	***p***
Would feel more secure in vehicle if system were permanently available to them (Sec. 1, 3a).	0.0% (0)	4.3% (1)	60.9% **(14)**	34.8% **(8)**	**22.39**	.0001
Would feel more secure if system were permanently available to other family members (Sec. 1, 3b).	0.0% (0)	4.3% (1)	52.2% (12)	43.5% (10)	**19.61**	.0002
Believe this system would be likely to help authorities deliver assistance (Sec. 1, 3c).	0.0% (0)	4.3% (1)	21.7% **(5)**	73.9% (17)	**31.78**	.0000

Note: Number in parentheses under percentage shows cell size.

Many responses to these questions significantly correlated with responses to other questions on the survey. For example, the level of agreement on responses to participants feeling more secure in their vehicles if this system were permanently available to them positively correlated with responses to feeling more secure if this system were permanently available to other members of their family, which in turn positively correlated with responses to reporting that the operator almost always attempted to identity and describe their location. Also, the level of agreement on responses to believing this system would be likely to help authorities deliver assistance when participants are in situations requiting police, medical, or roadside assistance positively correlated with responses to reporting that the operator almost always attempted to identify and describe their location, as well as feeling more secure in their vehicle if this system were permanently available to them and other members of their family. The complete set of survey questions which had responses that significantly correlated with responses to questions on security and safety, and the relevant statistics, appear in Appendix J.

When respondents were asked if they felt confident that the operator would be able to correctly identify their location if the system were used in an actual emergency (question 2f, Section l), significantly more respondents answered yes (78.3%) than no (21.7%) ($c^2 = 7.35$, $p = .0067$). Respondents who answered *no* commented that while the operators could identify their location most of the time, in some cases the operators were not accurate enough. Two respondents commented that the response center maps needed updating and expanding to include more counties. Two respondents commented that when they were under overpasses the operators were unable to update their locations. The survey also asked if respondents could think of circumstances under which this system would NOT be able to help them. Over half of respondents (52.2%) answered yes, while the remaining 47.8% answered *no*. Three respondents commented that if they had dead batteries or their engine lost power the system would be unable to help them Two respondents commented that the unit was not portable, so if they needed to evacuate the vehicle (e.g., vehicle on fire), the system would not work. Five respondents commented that if they were in remote locations, out of cellular coverage, or in regions where the response center maps were outdated, the system would be unable to help them

Overall, most respondents had a very favorable attitude toward the safety and/or security that this system could offer, although several identified situations in which the system would be unable to help them as it was intended.

3.5.1.4 Motorola User Perceptions of Reliability and Consistency

Users were also asked several questions relating to their perception of the reliability or consistency of the system. Answers to these questions were generally very positive. As shown in Table 3.5, nearly all respondents *(95.2%)* reported that they *almost always* heard the phone beep and begin dialing shortly after pressing the EMER, RA, or TA button. Most respondents (90.9%) reported that *only rarely* or *occasionally* were they disconnected when speaking with the response center operator, and 100% reported that they were *almost always* or *frequently* automatically reconnected. Performance data (See 2.4.4.2) confirmed that 81% of disconnected calls were automatically reconnected.* Chi square analyses showed significant differences in responses for these questions

*** This should not be confused with the issue of whether or not the operator recognized the reconnected call as a new or previous call. (See Section 2.4.5)**

with a majority of respondents *almost always* hearing the phone beep begin dialing; *rarely* or *occasionally being disconnected;* and when disconnected, *almost always being* reconnected.

Table 3.5 - Motorola User Perceptions of Reliability

Question	Rarely	Occasionally	Frequently	Almost Always	c2	*p*
Heard the phone beep and begin dialing (Sec. 1, 1a).	0.0% (0)	4.8% (1)	0.0% (0)	95.2% (20)	55.38	.0000
When speaking with the response center operator, was disconnected (Sec. 1, 1f).	40.9% (9)	50.0% (11)	9.1% (2)	0.0% (0)	15.45	.0015
When disconnected, was automatically reconnected (Sec. 1, 1g).	0.0% (0)	0.0% (0)	18.2% (4)	81.8% (18)	39.82	.0000

Note: Number in parentheses under percentage shows cell size.

As shown in Table 3.6, most respondents *agreed* or *strongly agreed* that: (1) the time the operator took to respond to the call was usually consistent (89.9%); (2) it seemed to take only a short time for the operator to respond (86.4%); and (3) the system was consistent in how it worked (82.6%). Chi square analyses showed significant differences in responses for these questions, with responses favoring the *agree* response.

Table 3.6 - Motorola User Perceptions of Consistency

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	c2	*p*
Felt that the time the operator took to respond to call was usually consistent (Sec. 1, 3n).	4.5% (1)	4.5% (1)	81.8% (18)	9.1% (2)	38.00	.0000
Felt that it seemed to take only a short time for the operator to respond (Sec. 1, 3o).	0.0% (0)	13.6% (3)	86.4% (19)	0.0% (0)	45.27	.0000
Felt the system was consistent in how it worked (Sec. 1, 3p).	0.0% (0)	17.4% (4)	65.2% (15)	17.4% (4)	21.70	.0001

Note: Number in parentheses under percentage shows cell size.

Many responses to these questions significantly correlated with responses to other survey questions. For example, the level of agreement on responses regarding the consistency of the time it took the operator to respond to participants' calls positively correlated with responses to reporting that the operator correctly described their location. The complete set of survey questions with responses that significantly correlated with responses to the reliability and consistency questions, and the relevant statistics, appear in Appendix J.

Overall, most respondents found the operation of the Motorola system quite reliable and consistent.

3.5.1.5 Additional Motorola User Perceptions of the Testing or System

Several other questions gathered feedback from the users on some of the testing procedures or on the system in general. Although in the User Group Deployment, the CSC Operators were not required or instructed to identify and describe to users their location, in practice the operators usually did. As shown in Table 3.7, 87% of respondents reported that the operator *almost always* attempted to identify and describe their location. The operators were apparently often successful: 60.9% of respondents reported that the operator *almost always* correctly described their location, while the remaining 39.1% of respondents indicated that the operator *frequently* correctly described their location. Chi square analyses showed significant differences in responses to these questions, with the responses favoring the positive end of the scale. In addition, responses to these two questions positively correlated ($r = .4517$, $p = .030$).

Table 3.7 - Additional Motorola User Perceptions of the Testing or System

Question	Rarely	Occasionally	Frequently	Almost Always	c2	*P*
Operator attempted to identify and describe location (Sec. 1, 1h).	0.0% (0)	4.3% (1)	8.7% (2)	87.0% (20)	47.43	.0000
Operator correctly described location (Sec. 1, 1i).	0.0% (0)	0.0% (0)	39.1% (9)	60.9% (14)	25.17	.0000

Note: Number in parentheses under percentage shows cell size.

Respondents were also asked if anything occurred during the tests that was unexpected. About half (54.5%) indicated that something did occur that was unexpected. Two respondents commented that the computer at the CSC went down several times. Two respondents indicated that sometimes the operator identified their location at places they had been at earlier in the day. One respondent indicated that once, after pressing the EMER button, the respondent was connected with an answering machine. Another indicated that the system often disconnected and reconnected before the operator actually answered.

When asked if they felt there were some features missing in the system, significantly more respondents answered no (97%) than yes (3%) ($c^2 = 12.57$, $p < .0l$). One respondent commented that a louder sound system would help with hearing. Another suggested a "send help regardless" button so that the response center could respond to a call when the person could not be in the vehicle. Finally, when asked if they believe this system could provide a service not encountered before, significantly more respondents answered yes (95%) than no (5%) ($c^2 = 18.18$, $p < .0l$).

3.5.1.6 Relationships Between Motorola Demographic Characteristics and Perceptions of the System

Respondents' demographic characteristics were used to determine if there were significant differences or relationships between demogmphic characteristics of the respondents and responses to other survey questions. The following demographic characteristics were analyzed: gender; age; income; occupation; average annual vehicle miles traveled; and the use or ownership of cellular phones, pagers, or home security systems. Because most respondents were in technical occupations, the survey responses of those in non-technical occupations were grouped together and compared to the responses of those in technical occupations. For the analyses of survey responses grouped by respondents' income, only income groups with more than three respondents were used ($30,000 to $39,999; $40,000 to $49,999, and $50,000 to $75,000). Differences in education could not be used for grouping purposes since most respondents indicated similar levels of education (community college or college educated).

See Appendix K for a complete list of significant differences and relevant statistics.

3.5.1.6.1 Motorola Significant Differences in Responses by Gender, Age, Income, and Occupation

When respondents were asked if they believe this system could provide a service not encountered before, males were significantly more likely than females to answer yes. Females, though, were significantly more likely than males to report that they found the microphone more effective than the handset and that they preferred using the microphone, while males were significantly more likely than females to indicate that they preferred using the handset. The age of respondents' significantly correlated with responses to two survey questions of interest. Age of respondents positively correlated with feeling more secure in their vehicle if this system were permanently available to them, and inversely correlated with finding the operator's voice at the response center easy to understand In addition, there was nearly statistical significance in age positively correlating with feeling more secure if this system were permanently available to other family members. The analyses indicated no significant differences in responses to any questions when respondents were grouped by income or occupation.

3.5.1.6.2 Motorola Significant Differences in Responses by Cellular Phone Ownership, Use of Pagers, and Home Security Service Subscription

When respondents were asked if there were any features missing in the system that they would have expected to be included, those who *did not* own cellular phones were more likely to answer yes. In addition, those who *did not* use pagers were more likely to (1) feel more secure if this system were permanently available to other members of their family, (2) believe this system would be likely to help authorities deliver assistance when they are in situations requiring police, medical, or roadside assistance, and (3) prefer to use the microphone. Those who *did* use pagers were more likely to agree that they preferred to use the handset. There were no significant differences in responses to questions by those with home security service subscriptions.

3.5.2 XYPOINT RESULTS

This section presents results of the surveys administered to volunteers using the XYPOINT system. The results below discuss user perceptions of ease of use; user perceptions of safety/security; user perceptions of reliability and consistency; additional user perceptions of the testing or system; and the relationships between demographic characteristics and perceptions of the system.

3.5.2.1 XYPOINT User Perceptions of Ease of Use

The XYPOINT survey included several questions that focused on respondents' perceptions of general ease of use of the device or system. With the exception of a few questions, answers to these questions were generally positive. When asked if they found the device easy to reach (question 1a, Section 1), significantly more respondents answered yes (67.9%) than *no* (32.1%) ($c^2 = 6.81$, $p < .01$). Similarly, when asked if they found the device easy to handle (question 1b, Section 1), 61.1% answered yes and 38.9% answered *no*. A number of respondents commented that the cords of the in-vehicle unit were too long and easily tangled; there was no convenient location to place the in-vehicle unit; the unit was inconvenient to store; the unit should be permanently mounted, and it was inconvenient to have to plug it in and attach the antenna on the roof.

As shown in Table 3.8, most respondents either *agreed or strongly agreed* that they found the device easy to use (84.9%%), the written-instructions easy to understand (90.7%), the device easy to set-up (82.6%), and the messages easy to read (81.1%) and easy to understand (94.0%). Most respondents *agreed* or strongly agreed (64.8%) that they knew what to expect when they operated *the* device, although *20.4% disagreed* and 14.8% *strongly disagreed. Chi* square analyses showed significant differences in responses to these questions, with responses favoring the *agree* response.

When asked if respondents were aware each time a new message appeared on the screen, 49% responded yes and 51% responded *no* (section 1, question lc). Many respondents commented that the messages changed too fast; there was no indication that they had received a new message; the message did not stay on the screen long enough; and they only saw the messages if they were looking at the device all the time. As shown in Table 3.8, most respondents *agreed* or *strongly agreed* that they: (1) sometimes missed seeing a new message appear on the screen (96.0%), (2) found the way the device beeped helpful (77.1%), but (3) felt it would be helpful if the device beeped every time a new message appeared on the screen (98.1%). Chi square analyses showed significant differences in responses for these results, with most respondents *agreeing or strongly agreeing.*

Further, most respondents felt confident in selecting which button to push, with 22.6% indicating that they *strongly agreed,* and 56.6% indicating that they *agreed. The results* of the user group deployment (see Section 2.2) indicate that for most scenarios, there was consistency in the button users pushed For 10 of 16 scenarios tested in the user group deployment, there was 90% or greater consistency in the button pushed by users (see Appendix I for detailed data). For 3 of 16 scenarios, there was 80-89% consistency; for 2 of 16 scenarios, there was 60-69% consistency; and for one scenario there was 5060% consistency.

Table 3.8 - XYPOINT User Perceptions of Ease of Use

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	c^2	p
Found device easy to use (Sec. 1, 2h).	0.0% (0)	15.1% (8)	60.4% (32)	24.5% (13)	41.87	.0000
Found the written instructions for using the device easy to understand (Sec. 1, 2f).	0.0% (0)	9.3% (5)	66.7% (36)	24.0% (13)	56.37	.0000
Found device easy to set-up (Sec. 1, 2e).	3.8% (2)	13.5% (7)	63.4% (33)	19.2% (10)	48.54	.0000
Found messages easy to read (Sec. 1, 2i).	1.9% (1)	17.0% (9)	60.4% (32)	20.7% (11)	39.60	.0000
Found messages easy to understand (Sec. 1, 2j).	0.0% (0)	6.0% (3)	64.0% (32)	30.0% (15)	50.64	.0000
Knew what to expect when operated device (Sec. 1, 2d).	14.8% (8)	20.4% (11)	55.5% (30)	9.3% (5)	28.22	.0000
Sometimes missed seeing a new message appear on screen (Sec. 1, 2o)	2.0% (1)	2.0% (1)	49.0% (25)	47.0% (24)	43.35	.0000
Found the way the device beeped helpful (Sec. 1, 2k).	2.1% (1)	20.8% (10)	50.0% (24)	27.1% (13)	22.50	.0001
Would be useful if device beeped every time a new message appeared (Sec. 1, 2n)	1.9% (1)	0.0% (0)	47.2% (25)	50.9% (27)	49.26	.0000
Felt confident in selecting which button to push (Sec. 1, 2g).	0.0% (0)	20.8% (11)	56.6% (30)	22.6% (12)	34.92	.0000

Note: Number in parentheses under percentage shows cell size.

Many responses to these questions significantly correlated with responses to the other survey questions. For example, the level of agreement to finding the device easy to use positively correlated with responses to finding the device easy to set-up, finding the written instructions for using the device easy to understand, and feeling confident in selecting which button to push for each scenario given. The level of agreement to finding the written instructions easy to understand positively correlated with responses to finding the device easy to set-up, which in turn correlated with responses to knowing what to expect when they operated the device. The level of agreement to finding the messages easy to read positively correlated with responses to finding the device easy to use. The level of agreement to finding the messages easy to understand positively correlated with responses to feeling confident in selecting which button to push, and finding the device easy to use. Finally, the level of agreement to sometimes not seeing new messages appear on the screen positively correlated with responses to agreeing that it would be helpful if the device beeped every time a new message appeared on the screen. The complete set of survey questions with responses that significantly correlated with responses to the ease of use questions, and the relevant statistics, appear in Appendix J.

Overall, most respondents reported that they found the device relatively easy to use and handle, although a number had difficulty with cords and the set-up, and even more had difficulty with the text message display, missing new messages either because the messages changed too quickly or they did not receive a clear enough indication from the system that a new message had arrived.

3.5.2.2 XYPOINT User Perceptions of Security/Safety

The questionnaire also included several questions regarding user perceptions of safety or security that this system could offer. Answers to these questions were more positive than negative. As shown in Table 3.9, 69.8% of respondents *agreed or strongly agreed* that they would feel more secure in their vehicle if this system were permanently available to them, while 30.2% *disagreed* or *strongly disagreed.* When respondents were asked if they would feel more secure if this system were permanently available to other members of their family, 74% *agreed or strongly agreed, while 26% disagreed or strongly disagreed.* When respondents were asked if they believe this system would be likely to help authorities deliver assistance when they are in situations requiring police, medical, or roadside assistance, 85% *agreed* or *strongly agreed, while* 15% *disagreed* or *strongly disagreed. Chi* square analyses showed significant differenccs in responses to these questions, with responses favoring the *agree* response.

Table 3.9 - XYPOINT User Perceptions of Security/Safety

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	c2	p
Would feel more secure in vehicle if system were permanently available (Sec. 1, 2a).	9.4% (5)	20.8% (11)	54.7% (29)	15.1% (8)	26.32	.OOOO
Would feel more secure if system were permanently available to other family members (Sec. 1, 2b).	8.0% (4)	18.0% (9)	54% (27)	20.0% (10)	24.08	.0000
Believe this system would be likely to help authorities deliver assistance (Sec. 1, 2c).	7.5% (4)	7.5% (4)	66.1% (35)	18.9% (10)	49.42	.0000

Note: Number in parentheses under percentage shows cell size.

Many responses to these questions significantly correlated with responses to other questions on the survey. For example, the level of agreement to believing this system would be likely to help authorities deliver assistance when participants are in situations requiring police, medical, or roadside assistance positively correlated with responses to feeling more secure if this system were available to other members of their family, and feeling more secure in their vehicle if this system were permanently available to them The complete set of survey questions with responses that significantly correlated with responses to the questions regarding security and safety, and the relevant statistics, appear in Appendix 3.

When respondents were asked if they felt confident that the operator would be able to correctly identify their location if the system were used in an actual emergency (question lg, Section l), slightly more respondents answered yes (53.2%) than *no* (46.8%). About 14 respondents (26%) commented that they did not get any feedback on their location from the operator. Several respondents commented that they did not feel confident that the operator would be able to locate them because the system was unreliable. When respondents were asked if they could think of any circumstances under which they felt this system would *not* be able to help them acquire emergency services (Section 1, question lh), significantly more respondents answered yes (69.6%) than *no (30.4%)* ($c^2 = 7.04$, $p < .0l$). About 10 respondents commented that coverage was not good in all situations or locations (e.g., in tunnels, in parking garages, outside of cellular areas, under overpasses, near power lines, when cellular phones are in use nearby). A number of respondents commented that the system was too unreliable. Some respondents may have been negatively impacted by the need to set-up the system, a feature of the test system that would not be part of a marketed system (e.g., the comment-- "If I were truly hijacked, I doubt my abductors would let me set up the machine").

Overall, a majority of respondents agreed that this system would provide some feelings of safety or security to them, although a majority identified situations in which the system would not be helpful In addition, a number of respondents expressed reliability concerns that negatively impacted their sense of security.

3.5.2.3 XYPOINT User Perceptions of Reliability and Consistency

Users were also asked several questions relating to their perception of the reliability or consistency of the system. Answers to these questions were mixed. Answers to questions regarding the device beeping regularly were generally positive, but answers to questions regarding the overall consistency of the device were generally negative. As shown in Table 3.10, most respondents *agreed or strongly agreed* that the device beeped three times when ready (75.6%), and that it beeped once after pushing a button (84.2%). Chi square analyses showed significant differences in responses to these questions, with responses favoring the *agree* response.

When asked if it seemed to take only a short time to receive a message from the response center, 51% *disagreed* or *strongly disagreed,* while *49% agreed* or *strongly agreed.* Although chi square analysis showed significant differences in responses to this question, the significance is probably due to the *strongly disagree, disagree,* and *agree* responses being chosen significantly more frequently than the *strongly agree* response (one respondent).

When asked if the time it took to receive an initial response was usually consistent, 70% of respondents *disagreed* or *strongly disagreed, while* 30% *agreed* or *strongly agreed.* When asked if they felt the system was consistent in how it worked, 7 1.1% of respondents *disagreed* or *strongly disagreed, while* 28.9% *agreed* or *strongly agreed.* Chi *square analyses* showed significant differences in responses to these questions (with responses favoring *the disagree* response), although the significant differences may be due to the extremely low response for *strongly agree* (one respondent).

Table 3.10 - XYPOINT User Perceptions of Reliability and Consistency

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	c^2	*p*
Device beeped three times when ready (Sec. 1, 21).	11.1% (5)	13.3% (6)	57.8% (26)	17.8% (8)	26.20	.0000
Device beeped once after pushing button (Sec. 2, 2m).	4.5% (2)	11.3% (5)	63.7% (28)	20.5% (9)	37.27	.0000
Seemed to take only a short time to receive message (Sec. 2, 2p).	22.4% (11)	28.6% (14)	47.0% (23)	2.0% (1)	20.14	.0002
The time it took to receive initial response was consistent (Sec. 2, 2q).	28.0% (14)	42.0% (21)	28.0% (14)	2.0% (1)	16.72	.0008
Feel the system was consistent in how it worked (Sec. 2, 2r).	25.0% (13)	46.1% (24)	27.0% (14)	1.9% (1)	20.46	.0001

Note: Number in parentheses under percentage shows cell size.

Many responses to these questions significantly correlated with responses to other survey questions. For example, the level of agreement on responses regarding the time it took to receive an initial response being consistent positively correlated with responses to feeling more secure if this system were permanently available to them and to other members of their family. Also, the level of agreement to feeling that the system was consistent in how it worked positively correlated with responses to knowing what to expect when they operated the device, and reporting that the time it took to receive a message was usually consistent. The complete set of survey questions with responses that significantly correlated with responses to the reliability and consistency questions, and the relevant statistics, appear in Appendix J.

Overall, based on responses to questions focusing on users perceptions of the reliability and consistency of the system, most users found that the system performed inconsistently.

3.5.2.4 Additional XYPOINT User Perceptions of the Testing or System

Several other questions focused on gathering feedback from the users on some of the testing procedures or on the system in general. When asked if anything occurred during the tests that was unexpected (Section 1, question 1d), more respondents answered yes (60.8%) than no (39.2%). A number of respondents commented that it was less reliable than they expected. Several respondents indicated that they received "odd" messages or symbols.

When asked if there were any features missing in the system that the respondent would have expected to be included (Section 1, question 1f), 53.1% of respondents answered *no*, while 46.9% of respondents answered yes. Several respondents commented that they needed better feedback from the system A few respondents (two to four) commented that: (1) they would have liked to have their location transmitted to them; (2) the screen should be illuminated for night use; (3) the device should have a cancel function; (4) the device should have a button to push in case of fire; (5) they would like more warning that a message was changing; and (6) they would like to have voice contact with the operator.

When asked if they believed that this system could provide a service that they have not encountered before (Section 1, question le), significantly more respondents answered yes (84.0%) than *no* (16.0%) ($c^2 = 23.12$, $p < .0l$). However, six respondents commented that a cellular phone would be more practical, reliable, or easier, and five respondents commented that it could provide a service they have not encountered if it was more reliable.

3.5.2.5 Relationships Between XYPOINT Demographic Characteristics and Perceptions of the System

Respondents' demographic characteristics were used to determine if there were significant differences or relationships between demographic characteristics of the respondents and responses to other survey questions. The following demographic characteristics were analyzed: gender; age; income; occupation; average annual vehicle miles traveled; and the use or ownership of cellular phones, pagers, or home security systems. Because most respondents were in technical occupations, the survey responses of those in non-technical occupations were grouped together and compared to the responses of those in technical occupations. For the analyses of responses by income, only income groups with more than three respondents were used ($30,000 to $39,999; $40,000 to $49,999; and $50,000 to $75,000). Differences in education could not be used for grouping purposes since most respondents indicated similar levels of education (community college or college educated).

See Appendix K for a complete list of significant differences and relevant statistics.

3.5.2.5.1 ***XYPOINT Significant Differences in Responses by Gender, Age, Income, and Occupation***

Analyses of responses by occupation indicated significant differences in responses to two questions. Respondents in non-technical occupations were significantly more likely to report that they believed this system would be likely to help authorities deliver assistance when in situations requiring police, medical, or roadside assistance, and that they found the device easy to use. The age of respondents inversely correlated with being able to think of situations under which respondents felt the system would *not* be able to help them acquire emergency services. In addition, there was nearly statistical significance in age inversely correlating with finding the messages on the device easy to understand Other analyses indicated no significant differences in responses to any questions by gender or income.

3.5.2.5.2 XYPOINT Significant Differences in Responses by Cellular Phone Ownership, Use of Pagers, and Home Security Service Subscription

Analyses of responses to *yes/no* by cellular phone ownership indicated significant differences in responses to one question of interest in each of these categories. When respondents were asked if they could think of any circumstances under which this system would *not* be able to help them acquire emergency services, respondents who *did not* own cellular phones and those whose homes were *not* protected by home security systems were significantly more likely to indicate that they could think of circumstances in which this system would *not* be able to help them acquire emergency services. Analyses of responses by pager use indicated that respondents who did *not* carry pagers with them were significantly more likely to indicate that they found the device easy to reach. Analyses of responses to other survey questions found no significant differences based on cellular phone ownership, pager use, or subscription to a home security system

3.6 DISCUSSION AND CONCLUSIONS

Before the usability analysis of each individual system is summarized, some general points need to be made. Usability testing was performed on the PuSHMe devices as they existed at the time of the operational test In other words, we evaluated what the systems were, not what they could be. (In fact, this usability analysis contains information that can help in future design modifications.) In each case (particularly XYPOINT), the current state of the device and the systems that supported it were far from market-ready. So, for example, many XYPOINT problems with consistency of performance were undoubtedly related to the newer CDPD network, which the XYPOINT system relied upon. During the operational test, this CDPD network was undergoing upgrades and improvements, and it is reasonable to expect that as improvements were made to the CDPD network, users' perceptions of the reliability of the system would improve as well.

On the other hand, it may be that the high level of education, income, and technological sophistication of the user population (not to mention their association with companies willing to contribute to the test) made them more receptive to using the technology and seeing benefits to that use. In a previous assessment of a technology-based dynamic ridesharing system, the evaluators found "a dichotomy between the desire to use the [system's] information and willingness to use the [system's] technology. Lower income employees were significantly more likely to use the information offered by the [ridesharing system] than were higher income employees; however, the lower income employees were also significantly less comfortable with various technologies."*

3.6.1 DISCUSSION AND CONCLUSIONS FOR THE MOTOROLA SYSTEM

Overall, the response of the users' to Motorola's in-vehicle emergency response system was very favorable in nearly every respect. The responses to questions focusing on user perceptions of general ease of use of the system indicate that in nearly every respect there is widespread satisfaction with the ease of use of the system. Regarding differences between the handset and the microphone, most respondents reported that they could hear the operator and vice versa using either, although the handset appeared to lead to clearer reception (particularly with the operator being able to hear the user). Most respondents, however, still preferred to use the microphone, which enabled hands-free communication. Most users found the device easy to use, reliable, and consistent. Finally, most respondents had a very favorable attitude toward the safety and/or

*** Haselkom, M., Spyridakis, J., Blumenthal, C., Michalak, S., Goble, B. and Gamer, M. (1995) *Bellevue Smart Traveler: Design, Demonstration, and Assessment,* Technical Report WA-RD 376.1, Washington State Transportation Center, p. xii.**

security that this system could offer, although several identified situations in which the system would be unable to help them as it was intended.

Slightly more than half of respondents did point out that the system did not always work as they expected (e.g., computers went down or the operator identified them at a location that they had been at earlier in the day). Several participants, however, expressed enthusiasm for the system For example, one respondent commented that there would be a market for a system like this, while another respondent commented on the reliability with which the operator described his/her exact location.

3.6.2 DISCUSSION AND CONCLUSIONS FOR THE XYPOINT SYSTEM

Overall, the response of the users to XYPOINT's in-vehicle emergency response was favorable in many respects and unfavorable in a few. Regarding perceptions of ease of use, most respondents reported that they found the device relatively easy to use and handle, although a number did comment that the cords tangled easily and that it was awkward to set-up the device in their vehicles. In addition, many respondents commented that they sometimes missed seeing new messages appear on the screen. Regarding feelings of safety or security, a majority of respondents agreed that this system would provide some feelings of safety or security to them, though this judgment was tempered by the fact that most users found that the system performed inconsistently.

This last point, that the system was not consistent in how it worked, was the most common issue respondents brought up. In addition, several respondents commented that they would have liked the system to have allowed them to speak with the operator.

The issue of usability is closely related to the issue of marketability; in fact our market study borrowed from this usability study to help design its analytical instruments. Chapter 4 takes this evaluation from concerns of use and acceptance of these systems to the related concerns of potential market for these systems.

CHAPTER 4. MARKET ANALYSIS

4.1 OBJECTIVES OF MARKET ANALYSIS

The goal of the market analysis was to evaluate market demand for in-vehicle emergency response systems (IVERS) such as those used in the PuSHMe operational test. Our objectives were to identify the conditions that characterize demand and assess the response of drivers in the Puget Sound region to the concept of an IVERS. The market analysis had two general purposes: (1) to provide guidance to the PuSHMe partners as they refine the designs of the systems that they will ultimately bring to market, and (2) to provide information to the government and other interested parties on likely market scenarios for these products.

4.2 STUDY DESCRIPTION

The study was comprised of three stages of data collection and analysis: (1) an instrument calibration stage, which helped in the development of instruments for the next stages by establishing a vocabulary that users (participants in the field test) and potential users (non-participants in the field test) employed to describe and evaluate solutions to emergency situations, (2) a marketability response measurement to assess how users and potential users view alternative IVERS, after users in the field have experienced the system, and (3) a market analysis survey of potential users to determine their knowledge of and need for IVERS.

A conjoint analysis was employed in order to provide an early indication of:

- Consumers' perceptions of alternative hypothetical IVERS.
- Trade-offs that different prospective consumers would make in choosing an IVERS.
- The effect of price on consumers' evaluation of characteristics of IVERS.

4.3 GENERAL OVERVIEW OF THE STUDY METHODOLOGY

This study was based on a user value model. A user value model serves as a framework for determining how consumers value alternative mayday systems. This user value model is not to be confused with usability (as discussed in Chapter 3), which instead focuses on ease of use and the user's acceptance and perception of the design features of the system.

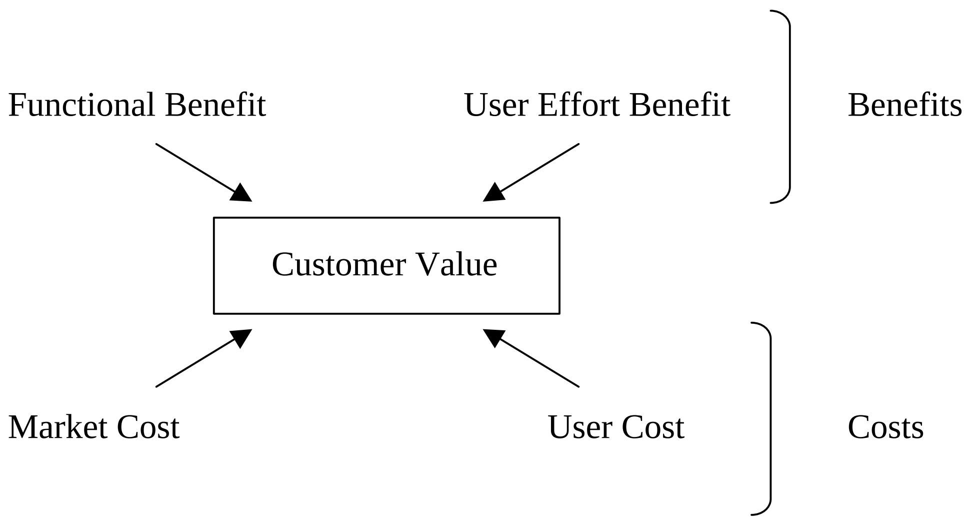
The value consumers derive from an IVERS will depend upon four categories of determinants. These are (as pictured below):

1. Functional benefits, which are directly related to engineering specifications. For example, one device may have voice communication while another relies on text messages.

2. User effort benefits, which depend on the ability and desire of the user to adapt a specific mayday system configuration to his or her specific needs. This is related to usability as described in Chapter 3.

3. User costs, which are the explicit and implicit costs people expect to incur in using alternative configurations. These costs depend on the user's current technology (this is described further below).

4. Market costs, which depend on prices and tariffs to the extent that they may be present.



The concept of user costs requires further explanation. User costs depend on the user's current technology for achieving the functionality provided by the new product. Functionality is determined and user costs are incurred when a product is used, that is, in the context of a consumption activity. By user's technology, we refer to anything the user can exploit in accomplishing the objectives of a specific consumption activity. Examples of consumption activities would be commuting from home to the workplace, going shopping, going skiing, and so forth. Users differ with respect to the technologies that they can exploit in pursuit of these consumption activities: some user's can exploit a public transportation system, some a bicycle, and others a private car. Even among those who use a car, technologies vary. For example, the user who commutes in a 1972 Volkswagen Beetle uses a decidedly different technology than someone who uses a 1995 Lexus. Additionally, users differ in terms of their views of consumption activities. For instance, the person who uses the VW Beetle to commute may view the consumption activity merely as transportation between home and the workplace; whereas the Lexus owner may bundle listening to classical music on a CD player into the daily commute because the ambiance of her car offers an appropriate listening environment. User costs are a mixture of out-of-pocket and significant implicit costs associated with gaining access to the service, adjusting to the service, learning how to exploit the service capabilities, and so forth.

Our market study attempted to discover the values that actual users and potential users attribute to alternative hypothetical IVERS. These valuations are conditioned on (a) a plausible consumption scenario, and (b) a set of relevant alternatives. Each alternative hypothetical IVERS was described in terms of the four categories of user values. The purpose of the analysis was to understand how individuals trade-off the four categories of user value in making a market assessment of a hypothetical mayday system and relevant alternatives.

4.4 METHODOLOGY

The methodology used for the IVERS market analysis consisted of three data collection stages. Stage 1 occurred before the operational field test, Stage 2 occurred during and after the field test, and Stage 3 occurred after the field test.

Stage 1 of the data collection was primarily an exercise to aid in structuring the Stage 2 market response measurement instrument. Stage 2 of the data collection involved users and potential users thinking of a mayday system in terms of broadly defined alternatives. Stage 3 of the data collection consisted of surveying potential users to determine their knowledge of emergency response systems, educate them if necessary, and assess their perceived need of such a system.

4.4.1 STAGE 1 METHODOLOGY

Prior to the PuSHMe field test, a sample of volunteers scheduled to participate in the user group deployment (before they had been introduced to the system) and non-participants were surveyed in order to: (a) communicate to them the concept of an IVERS, and (b) establish a framework for describing to them alternative IVERS. This pre-test was primarily to help determine how to structure the instrument for Stage 2. Appendix L includes a sample of the survey used in Stage 1.

Concept descriptions were prepared and given to respondents to communicate in simple terms the essential characteristics of an IVERS. Respondents were asked to articulate a set of relevant alternatives and to compare alternatives on dimensions that they considered to be important. Comparative dimensions were provided by the respondents and used to construct a product design matrix (see Appendix M, Figure la).

In addition, two scenarios were presented to respondents, each depicting a hypothetical breakdown under different conditions related to the responsiveness of conventional emergency road services. These scenarios provided a context for establishing some general evaluative criteria that respondents would apply in comparing conventional emergency road services and novel mayday services.

4.4.2 STAGE 2 METHODOLOGY

How potential customers understand and evaluate an in-vehicle emergency response system is tantamount to defining the market for these systems. These perceptions are partly revealed by the criteria people use in evaluation. To help us understand how people evaluate IVERS, we imposed a simple model of "customer value" based on that described in 4.3. We used this model to help us organize the interacting factors of an IVERS that influence an individual's evaluations. Based on these factors plus the information that was gathered in Stage 1, we constructed a paper and pencil exercise with different consumption scenarios. Under these scenarios, respondents provided feedback concerning their preferences for specific combinations of emergency response services.

The methodology employed in Stage 2 is known as profile analysis, which is a special form of conjoint analysis. The respondents were each given 18 cards with various combinations of system characteristics derived from the product design matrix (e.g. Appendix M, Figure 2). They were asked to rank order their preferences for the different emergency response systems described on the cards under two different cost scenarios. Demographic data was obtained from each respondent (when not available from other sources) and correlated with the preferences. This data collection provided the information that permitted calculation of marginal valuations of the critical dimensions of emergency response systems.

The sample for Stage 2 included both participants in the field test and non-participants. The analysis was conducted during and after the user group deployment,

4.4.3 STAGE 3 METHODOLOGY

Stage 3 consisted of a survey designed to evaluate non-participants' knowledge of emergency response systems, provide a description of a generic system to them if necessary, and determine their need for such a system (See Appendix N for a sample Stage 3 survey.) Demographic data were also obtained from each respondent. The goal was to gain a general sense of market awareness and need distinct from the operational test.

The Stage 3 survey was administered after the user group deployment. The sample was drawn from individuals living in the Puget Sound area who did not take part in the user group deployment (non-participants). Volunteers were gathered at three separate locations to maximize the diversity of the sample.

4.5 DATA ANALYSIS

4.5.1 STAGE 1 DATA ANALYSIS

The analysis for Stage 1 consisted of evaluating the survey results in order to construct a vocabulary to be used in the data collection tasks of subsequent stages. Stage 1 data included a listing of concerns or perceived costs that respondents cited in resolving two hypothetical driving incidents. Surveys were analyzed to determine the critical dimensions used by respondents to evaluate alternative lVERS. Additionally, the analysis established a set of emergency response alternatives that future users and potential users consider to be relevant.

4.5.2 STAGE 2 DATA ANALYSIS

For Stage 2 we analyzed the ranked data gathered from the conjoint analysis exercise. These data consisted of participants' ordering of alternative IVERS as described on profile cards. Following is a brief discussion of three issues that shaped this analysis: (1) Pricing Effects, (2) Grouping within the Sample, and (3) Estimation of Preferences. This is followed by a more detailed description of the overall analysis procedure.

4.5.2.1 Price Effects

Because the concept of an in-vehicle emergency response system is a relatively new one, we were concerned that drivers would have some difficulty in processing information describing alternative IVERS. It is well documented that when someone is asked to make a judgment or to express a preference for *unfamiliar* alternatives (such as IVERS), he or she is likely to use perceived price as a proxy or indicator of quality (i.e. if it costs more, it must be better). This means that price not only (1) serves to allocate the private resources of individuals, it also (2) conveys information about the imputed intrinsic properties of unfamiliar alternatives.

In order to understand the evaluations that individuals made of alternative IVERS configurations, we had to devise a means to isolate these two roles of price (allocative and informational). This was accomplished by having the respondents rank alternative systems under two different cost scenarios: (1) where they were to receive the system at no direct cost (purchase or usage) to themselves, and (2) where they would have to pay the listed purchase cost and usage fee. Under scenario 1, the purchase cost and usage fee were also listed on the cards, but respondents were told they would not have to pay for these directly. In this case, any price effect is informational, that is, price is used as a proxy for unknown characteristics. The ranking we conducted under scenario 1 is referred to as an unconstrained preference ranking -- i.e. without a budget restriction. The ranking we conducted under scenario 2 is referred to as a constrained preference ranking -- i.e. with a budget restriction. This case is closer to the choices people make in the market, where the informational and allocative roles of price are confounded.

4.5.2.2 Grouping within the Sample

A second issue concerned how we would use the different individual responses to alternative IVERS configurations to gain an understanding of how these evaluations represented distinct groupings of the consumers within the sample. To do this we employed a technique known as *Q-factor analysis.* Our basic goal was to group the individuals by similarity of responses. However, we also hypothesized that those individuals with some personal experience with IVERS were likely to have different evaluations than those individuals without any experience (i.e. participants vs. non-participants). Therefore, in performing the Q-factor analysis we first grouped by participation and then distinguished individuals from each other in terms of their preference structures. In so doing, we constructed clusters, or “representative consumers,” where each cluster is assembled on the basis of its *preference consistency.* In the end, we were pleased to find a strong relationship between general preference and characteristics of an IVERS.

4.5.2.3 Estimation of Preferences

Estimation of preferences was performed using regression analysis to capture multiple correlations of several indicators (i.e. system characteristics provided on the profile cards) with some overall preference variable (i.e. preference rankings as provided by the respondents). Two important issues arise in performing the estimation of preferences at both the aggregate and disaggregate levels -

1. Whether to represent variables as continuous or discrete.
2. Whether to use linear or nonlinear functional forms.

To the extent possible, we attempted to interpret the results using continuous representations of the independent variables. Additionally, we treated the ordinal preference data as continuous for estimation purposes. While this is an approximation of factors that are designed to vary on discrete levels, it allowed us to conserve degrees of freedom

In our handling of functional forms, we were faced with two possible ways that individuals might process information in choosing among alternative IVERS configurations. Individuals might link choices in reaching an overall preference judgment by balancing factors that they evaluate at low levels with factors that they evaluate at high levels. In this case a *linear additive* or *compensatory* model would be most appropriate For example, if factor A is evaluated at 1 and factor B is evaluated at 0, a linear-additive (compensatory) model would result in an overall preference evaluation of the two factors together as : 1+0=1. Alternatively, individuals might treat each ranking as an individual decision and not employ a compensatory logic in combining information. *In this* case, a *multiplicative* model would be most appropriate Using the same hypothetical example, the overall preference would be : 1x0=0. In this case, overall preference would be zero no matter how favorable (high) the evaluation of factor A so long as factor B is evaluated at 0.

For purposes of summarizing the results we choose that functional form in each case which displayed the best classical statistical properties. Both functional forms were estimated for all representative consumers.

4.5.2.4 Procedure

Given the above considerations, we adopted the following procedure for analysis of Stage 2 data.

Two preference orderings of 18 IVERS profiles, each a distinct combination of system characteristics, were obtained for each respondent, one under the constraint of a budget and the other with no budget constraint. The individuals were grouped into representative consumers based on the similarity of their preferences. Diagnostic tests on the goodness-of-grouping were performed, and average background traits were determined from the demographic information. Finally, the allocative effect of price was isolated by estimating a difference equation of constrained less unconstrained preferences at both aggregate (total sample) and disaggregate (groups) levels.

The ranked data were used to determine the individual utility functions for each respondent. Two different forms of the design matrix were used in creating the utility functions, one as shown in Appendix M, Figure lb, and one with dummy variables for the product design variable, leading to two functions for each budget scenario (i.e. four preference models for each user). The data from the first 16 cards only was used for estimation, with the last two card rankings used to assess the accuracy of the model. The data were reranked out of 16, then linear additive and multiplicative regressions models were conducted, and these models were used to predict the final two rankings. The rankings from the regression models were correlated (two tailed Pearson rank correlation) with the original rankings for each respondent's four models.

The strength of the correlations determined the groupings based on a two stage grouping process. First respondents were split into the natural categories of participants in the field test and non-participants and new regression models were formed for these subgroups. Then these groups were broken down further by constrained preference using a Q-factor analysis. Q-factor analysis identifies a relatively small number of factors (corresponding to distinct groups of individuals) that can be used to represent relationships among sets of many interrelated variables (in this case the card rankings). It forms linear combinations of these variables, based on the variable correlation matrix and selects as few factors as possible to explain the observed correlations. For this part of the analysis we chose the transformation method Varimax with oblique rotation and principal components analysis as the factor extraction method (in this method the first factor is the combination that accounts for the largest variance in the sample). This classified the participants and non-participants into groups which would take into account most of the variance in the two samples.

Consistency tests were then done to illustrate the advantages of the chosen grouping. For the consistency test, a dominance matrix (Gautschi & Rao, 1987) was calculated for each individual, then within each subgroup the matrices were summed to form subdominance matrices. The consistency score, the sum of all the non-zero values of the sub-dominance matrix divided by the number of values and group size, was calculated for each subgroup and compared to the score of the aggregate consumer. When this goodness-of-grouping score exceeded a certain threshold, the subgroup in question qualified as a representative consumer group.

Common characteristics of the aggregate and each subgroup were determined from the demographic and other survey information using a discriminant analysis. These characteristics can help us infer some motivation for the preferences of each group. **A** table of some of the more relevant characteristics, such as age, sex, income, and occupation, appears in Appendix Q.

4.5.3 STAGE 3 : SURVEYS OF NON FIELD TEST PARTICIPANTS

During Stage 3 we surveyed 71 non-participant individuals. These surveys were analyzed to assess trends in the general population as to awareness of IVERS, perceived need for IVERS, and desire for legislation requiring vehicles be equipped with emergency response systems. ln addition a correlation was done on the number of incidents respondents were involved in versus their evaluation of their need for an IVERS.

4.6 RESULTS

4.6.1 STAGE 1 RESULTS

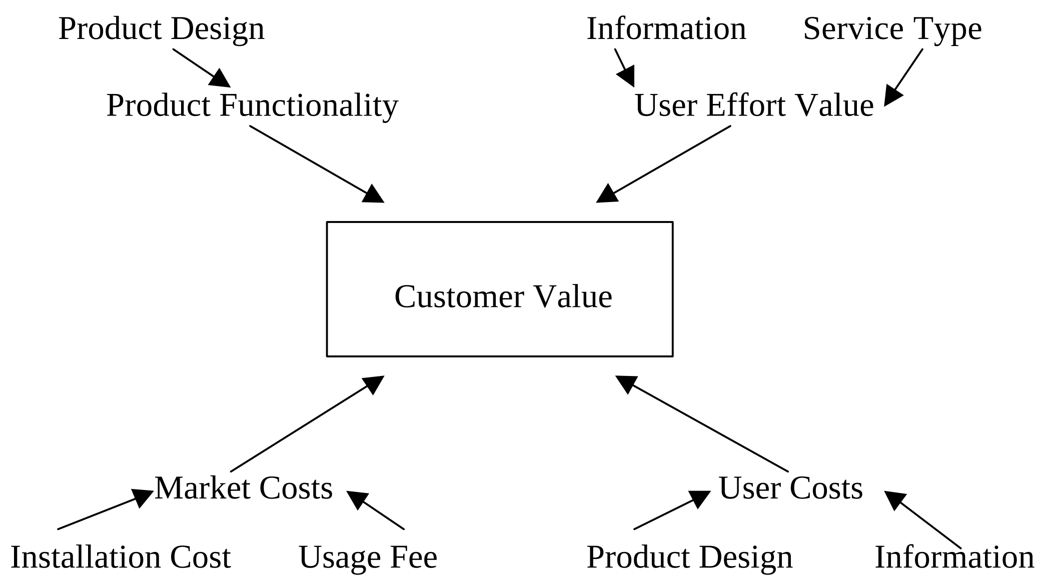
During Stage 1 we surveyed 19 individuals from the user population and 11 non-participants drawn from other Puget Sound sources. The results from these 30 surveys were used to develop the market assessment instrument required for Stage 2 as described below.

Analysis of the Stage 1 surveys resulted in the definition of five factors of influence as listed in the table below :

Factors	Levels
Product Design	(1) installation, (2) activation, (3) communication, (4) panel
Service Type	(1) services included or (2) no services
Information	(1) basic or (2) extensive
Purchase Cost	(1) $0, (2) $500, (3) $1000, (4) $1500
Usage Fee	(1) $0.00, (2) $0.50, (3) $2.00, (4) $5.00 (all per minute)

Product design covers various “technical” functions--how the product is installed (e.g. professionally vs. self), how communication is established (e.g. button push vs. voice activated), the nature of the communication (e.g. voice vs. text only), and panel features (e.g. lighted vs. not lighted). Service Type covers the extent of services provided (e.g. only respond to emergencies vs. not only respond to emergencies but also provide towing). Information covers the extent of information provided (e.g. only acknowledgment vs. acknowledgment plus directions to the nearest hospital, police station, etc.). Purchase cost means the initial cost of the device plus installation. Usage fee means the marginal price of using the system

Taken together, these five factors at their various levels permitted us to describe a wide variety of alternative IVERS offerings -- real and imagined. The five factors relate to the customer value model (as described in Section 4.3) in the following manner:



Note that a single factor can impact the model at more than one point. For example, product design factors will impact both functionality and user cost.

Using a fractional factorial design, we were able to construct 16 distinct alternatives as combinations of the five factors. These 16 alternatives, plus two additional alternative combinations that we added for predictive testing, became the 18 profile cards used in the conjoint analysis exercise. The profiles describe, in general terms, alternative IVERS configurations, as well as other market alternatives (e.g. cell phones) and non-market alternatives (i.e. forms of self-production). The 18 profiles and sample cards are presented in Appendix M.

4.6.2 STAGE 2 RESULTS

During Stage 2, 65 individuals were queried, 46 who participated in the field test (13 who used the Motorola system and 33 who used the XYPOINT system), and 19 potential users who did not participate in the field test

Q-factor analysis resulted in seven groups accounting for 94% of the variance for the participants, and four groups accounting for 88% of the variance for the non-participants. The consistency score of the aggregate was .7554 (out of a possible l.0), and the two subgroups of participants and non-participants had consistency scores of .768 and .7447 respectively. Q-factor analysis produced groups with consistency scores around .9 (see Appendix 0), a sizable improvement. These groupings were deemed acceptable and the regression models were rerun on these new groups. In all cases, we found estimating the models treating the variables as continuous to be satisfactory. Thus, all results apply to specifications with continuous, rather than discrete, variables.

The following summarizes the regression model's statistics and behavior both for the aggregate and for the various groupings of respondents. The results of estimating people's preferences were highly significant in all but one case, using both linear and non-linear specifications. The factors with the most significant influence on customer value are listed in the tables. The *beta weights are* a measure of the importance of each factor, with significance level *p* ($p \leq .15$ for the factor to be included), from either the linear additive or multiplicative regression model as noted. The most preferred alternative under each cost scenario was determined for the aggregate and for the larger subsequent groupings (n>5). The trends of the smaller groups are not clear cut, and probably not statistically significant, See Appendix R, which has the frequency each card was ranked first, second, or third under each cost scenario for all groupings. This illustrates obvious preference trends for the larger groups which are successfully predicted by the model. The smaller groups are formed based on similarity of preferences among group members, but it is more difficult to make generalizations. The predictive ability of the model for each subgroup was measured by comparing the rankings of the two test cards against the model's predictions. Appendix P contains a table of the resulting standard deviations for each cost scenario.

Unless otherwise noted, the results presented in this section indicate the influence of five factors (product design, service type, information, purchase cost, and usage fee) as presented in Section 4.6.1 within a realistic market context where choice is constrained by budget considerations.

4.6.2.1 AGGREGATE (n=65)

These results are for the entire sample, both participants and non-participants. A linear additive regression model was used.

Most significant influences on customer value:

|beta weight|	factor	p	
.443	purchase cost	.001	higher cost --> lower preference
.211	product design	.001	higher design --> greater preference
.07	usage fee	.05	higher cost --> lower preference
.039	service type	.11	higher level --> greater preference

For the aggregate, as expected, the higher the cost and usage fee the less the alternative was preferred, while the greater the functionality and service level the more the alternative was preferred. The extensiveness of information provided had no influence on preference in the aggregate. The highest rated alternative for the aggregate under a constrained budget scenario was a system with high end design, no services included, basic information, and no usage fee or purchase cost (the lowest cost model). Without a budget restriction the preference was for a more elaborate system, high end design, services included, extensive information, the highest usage fee, and a $500 purchase cost (the top of the line model, except for purchase cost). These results were consistent with either a linear additive or multiplicative regression model.

4.6.2.2 PARTICIPANTS VS NON-PARTICIPANTS

Participants (linear model) (n=46)

|beta weight|	factor	p	
.44	purchase cost	.0001	higher **cost --> lower** preference
.19	product design	.0001	higher design --> greater preference
.05	information	.11	higher level --> greater preference

Non-participants (linear model) (n=l9)

beta weight	factor	p	
.46	purchase cost	.0001	higher cost --> lower preference
.25	product design	.0001	higher design --> greater preference
.09	service type	.07	higher level --> greater preference

The simple disaggregation separating those who participated in the field test ("participants") from those who did not ("non-participants") produces some valuable information. The two groups are most sensitive to the same factor -- purchase cost -- and each group is sensitive to design features (participants slightly less so). Participants, however, value product information, whereas non-participants value service. Both groups are more sensitive to product cost than usage fees.

The highest rated alternatives were the same for the participant and non-participant subgroups as for the aggregate, the lowest cost model under a budget constraint and the model with top of the line features when cost wasn't a consideration. Again the results were consistent under either a linear additive or nonlinear regression model.

4.6.2.3 PARTICIPANT SUBGROUPS

The following subgroups represent a significant percentage of the sample population and/or are illustrative of a likely market segment. For smaller groups (n<5), we provide the table and description but have left out the top preferences.

Group 1 (linear model) (n=28)

| beta weight |	factor	p	
.57	purchase cost	.0001	higher cost --> lower preference
.09	product design	.0001	higher design --> greater preference
.07	service type	.01	higher level --> greater preference

This group represented 61% of the participants in the sample. It is a group that is much like the aggregate except it is less sensitive to usage fees, The highest rated alternatives for this subgroup were the same as for the aggregate, the lowest cost model under a budget constraint and the model with top of the line features when cost wasn't a consideration. A linear additive regression model was used in this analysis.

Group 2 (linear model) (n=5)

|beta weight|	factor	p	
.51	product design	.0001	higher design --> greater preference

This group represented 10.9% of the participants in the sample. For this group, the difference equation is insignificant. This means that there is no confounding of the price effect for this group -- their evaluations of IVERS alternatives are not influenced by a budget constraint. They are, however, positively influenced by design features. Not surprisingly, this group was the oldest, had the highest income, and traveled the most miles. A linear additive regression model was used.

Group 3 **(nonlinear model) (n=2)**

|beta weight|	factor	P	
.57	purchase cost	.001	higher cost --> lower preference
.26	product design	.15	higher design --> greater preference

This group represented 4.4% of the participants in the sample. It is a group that values purchase cost, but is not at all sensitive to usage fees. There is some indication that it is strongly influenced by design features as well. People in this group tend to be willing to pay a bit more for a better system. A nonlinear regression model was used.

Group 4 **(linear model) (n=5)**

|beta weight|	factor	p	
.72	product design	.0001	higher design --> greater preference
.27	purchase cost	.05	higher cost --> lower preference
.10	information	.14	higher level --> greater preference

This group represented 10.9% of the participants in the sample. It is a group that is influenced by purchase cost, with a strong preference for high product design and a desire for extensive information. Participants in this group are not at all sensitive to usage fees. **A linear additive** regression model was used.

Group 5 (linear model) (n=2)

| beta weight |	factor	p	
.64	service type	.001	higher level --> greater preference

This group represented 4.4% of the participants in the sample. It is a group that is not influenced by either purchase cost or usage fees, nor is it influenced by product design. This group is most sensitive to the presence of services. A linear additive regression model was used.

Group 6 (linear model) <<weak fit>> (n=2)

This group represented 4.4% of the participants in the sample. While none of the factor influences was statistically significant, there was some indication that this group is sensitive to usage fee - the higher the usage fee, the lower the preference - , and information - the higher the information level, the higher the preference. A linear additive regression model was used.

Group 7 (linear model) (n=2)

| beta weight |	factor	p	
.53	usage fee	.001	higher cost --> lower preference
.36	purchase cost	.05	higher level --> lower preference

This group represented 4.4% of the participants in the sample. A price sensitive group, it is most influenced by the usage fee and purchase cost None of the other factors is significant. A linear additive regression model was used.

4.6.2.4 NON-PARTICIPANT SUB GROUPS

Group 1 (nonlinear model) (n=9)

| beta weight |	factor	p	
.62	purchase cost	.0001	higher cost --> lower preference
.18	usage fee	.01	higher cost --> lower preference
.14	service type	.01	higher level --> greater preference

This group represented 47.4% of the non-participants in the sample. It is a group that is strongly influenced by both market cost factors with information and product design being the least important factors, similar to participant subgroup 7. This group also values service. The highest rated alternative under a constrained budget is a system with low end design, professional installation, services included, basic information, a low usage fee, and no purchase cost. With no budget constraint, the preferred system was the model with top of the line features and a purchase cost of $500. A nonlinear regression model was used.

Group 2 (nonlinear model) (n=4)

| beta weight |	factor	p	
.40	product design	.001	higher level --> greater preference
.20	purchase cost	.10	higher cost --> lower preference

This group represented 21.1% of the non-participants in the sample. It is a group that has no **desire** for extra information or services, but a strong desire for product design and low purchase cost. It is not sensitive to usage fees. A nonlinear regression model was used.

Group 3 (nonlinear model) (n=5)

| beta weight |	factor	p	
.69	product design	.0001	higher level --> greater preference
.18	**purchase cost**	**.15**	**higher cost --> lower preference**

This group represented 26.3% of the non-participants in the sample. It is similar to group 2, but even more sensitive to product design. Like group 2, there is no preference for extra information or services and no sensitivity to usage fees, but it is sensitive to purchase cost. A nonlinear regression model was used.

Group 4 (nonlinear model) (n=l)

| beta weight |	factor	p	
.64	purchase cost	.01	higher cost —> lower preference
.53	information	,05	higher level —> greater preference

This "group" (individual) represented 5.3% of the non-participants in the sample. It is a group that is most strongly influenced by purchase price, but also has a strong preference for extensive information. Neither product design nor service influences this group significantly. A nonlinear regression model was used.

4.6.3 STAGE 3 RESULTS

The Stage 3 sample was 71 individuals drawn from the Puget Sound area who did not take part in the user group deployment (non-participants).

The results from the Stage 3 analysis are summarized in the following charts. As shown in Figure 4.1, the respondents' previous knowledge of IVERS was low (72% had never heard of an emergency response device and an additional 7% were unsure).

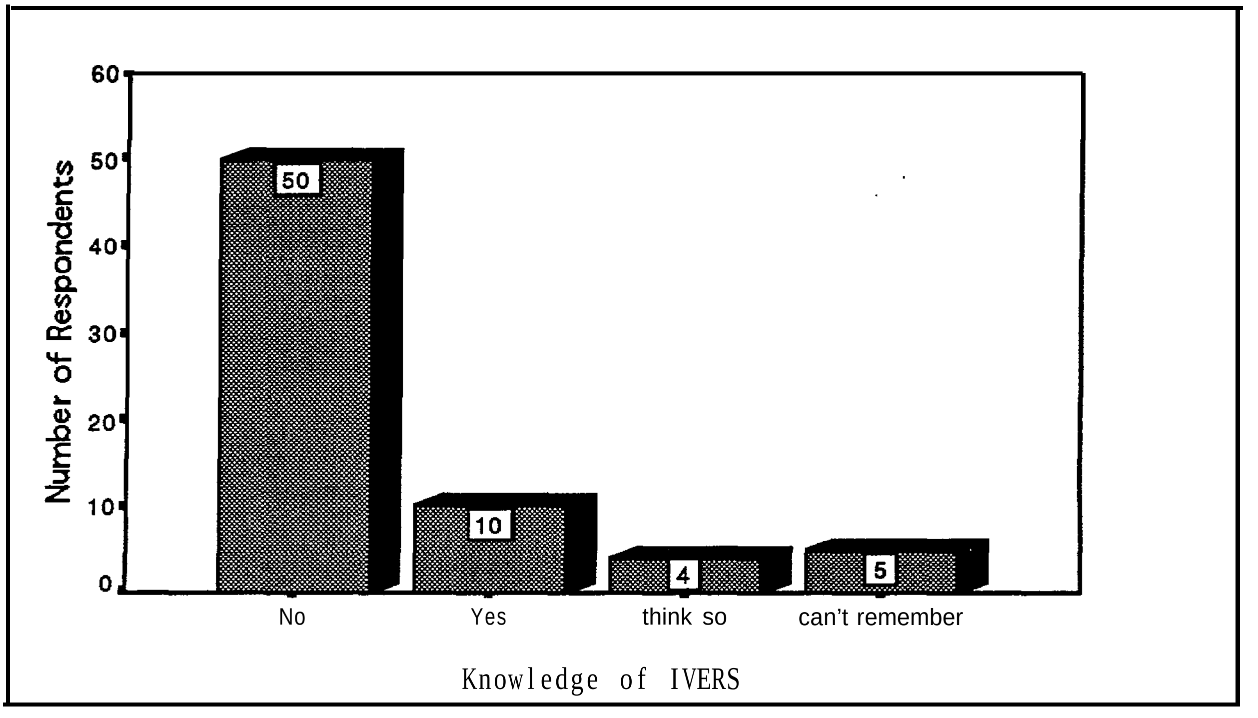


Figure 4.1- Stage 3 Survey Results, Previous Knowledge of IVERS

As illustrated in Figures 4.2 and 4.3, the respondents felt a neutral to weak need for an in-vehicle response device, and felt that there should probably not be legislation dictating the use of such devices. Only 15% of respondents had a strong or very strong need for an emergency response device, 37% were neutral and the rest (48%) had a weak or very weak need. 23% of respondents felt positively about legislation requiring vehicles to be equipped with emergency response devices, 27% were neutral, and the remaining 48% were opposed to the idea of such legislation. There was, however, a correlation between incidences of being lost and needing directions in the past year and need for an emergency response device, with correlation coefficient .23 and a significance level of .058 .

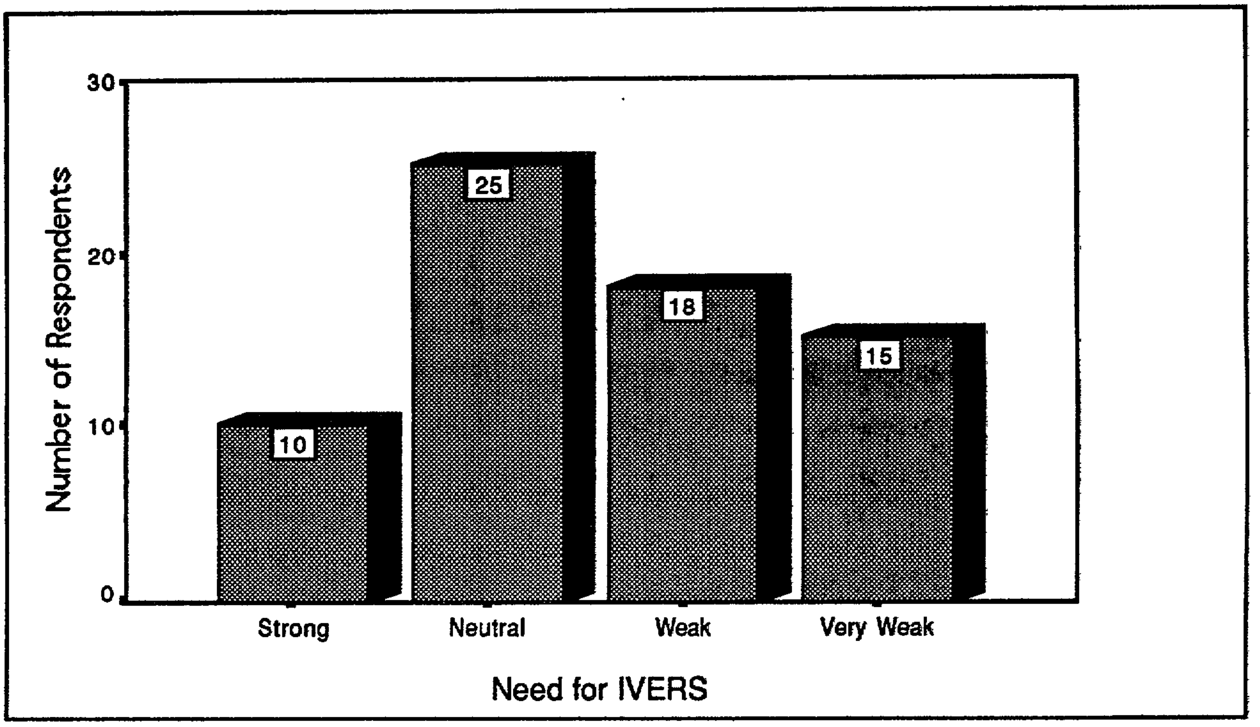


Figure 4.2 - Stage 3 Survey Results, Perceived Need for IVERS

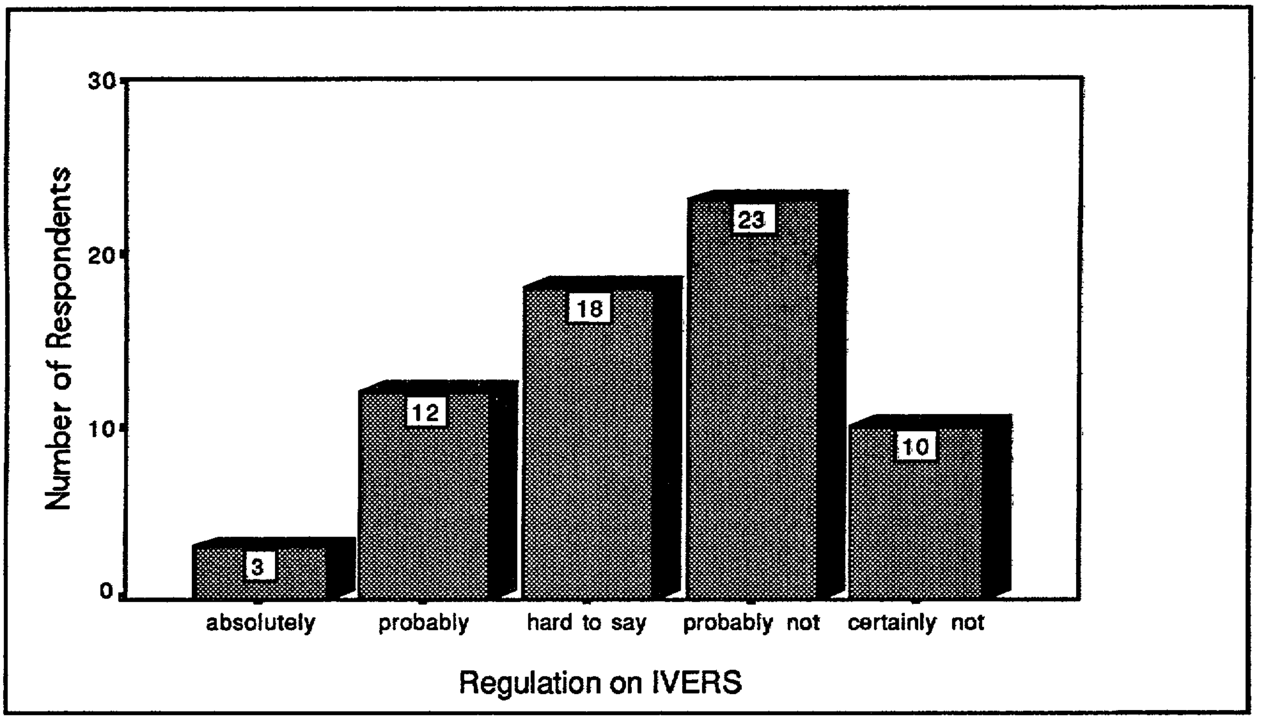


Figure 4.3 - Stage 3 Survey Results, Desire for Regulation of IVERS

4.7 CONCLUSIONS

Analysis of the aggregate revealed that cost is the key factor in the marketability of an IVERS, with purchase cost having far more impact than usage cost. This is not to say that people do not consider function--they do, but they do not generally want to trade higher cost for additional functionality. This is most graphically demonstrated by the differences in the system of choice under constrained and unconstrained conditions, with people desiring functionality in a non-constrained scenario, but sacrificing functionality for low purchase cost in a constrained scenario.

By splitting the aggregate sample into a group of those who had use experience and another group of those who had no use experience, we discovered that the secondary influences of service type and information differed across these groups. In general, use of an IVERS fostered a sensitivity to the value of additional information (e.g. directions), while non-users were more sensitive to service features (e.g. towing). This implies that initial marketing to encourage adoption of these systems might include bundled related services, while later efforts to maintain market share might focus on enhanced information services.

Further disaggregation of the participant and non-participant groups revealed additional insights. It is significant that the influences of purchase cost, usage fees, product design, information, and service levels are not at all consistent across the subgroups. This simple result indicates the critical importance to producers of IVERS of carefully segmenting the markets that they intend to supply. Imagine, for example, a market strategy for participant group 2 which was not at all sensitive to price, highly sensitive to functionality and consisted of older, richer, well traveled males; versus a market strategy for non-participant group 2 which wanted functionality at low cost and consisted of relatively low income (mostly) women who traveled less than one tenth the vehicle miles of their participant counterpart.

In only four subgroups (two participant, two non-participant) was product functionality the dominant influence on peoples' preferences for alternative IVERS. More consistently, purchase price of an IVERS alternative is the more important influence on peoples' preferences. Usage fees are significant influences in only two cases (dominant in only one case), and not at all influential in determining preferences in most cases. This implies a clear pricing strategy favoring usage fees for most market segments. There are no systematic patterns among subgroups in the influences of service or information.

The models were all subjected to some predictive testing and performed reasonably well. Figures 1 and 2 in Appendix P present the standard deviations in predicting the rankings of the 17th and 18th profiles (the test profiles). Generally, standard deviations are lower (thus, predictive accuracy is higher) for disaggregate analyses. The predictive results, thus, reinforce the conclusion that the population should be carefully segmented when designing and selling IVERS to the market.

Several interesting generalizations may bc drawn from the data analysis, especially the importance of segmenting the population of users and potential users of IVERS.It is not surprising that in almost all cases a linear compensatory form applies best to respondents with some use experience with IVERS. Since they have a 'concrete' understanding of the system, they are able to compensate for factors they don't particularly value by factors they value highly. This suggests that the Stage 1 calibration successfully yielded descriptions of alternative systems that experienced users could readily evaluate. Non-users, being less familiar with the concept of IVERS, were all more likely to evaluate alternative systems by fixing their attention on a small number of factors at the exclusion of others. This is illuminated by the fact that the non-linear non-compensatory form applies in all non-user subgroups.

Finally, as the Stage 3 results show, there are indications of palpable skepticism about the economic value of IVERS among randomly chosen people in the greater Seattle metropolitan area. Respondents indicated an aversion to government subsidy of IVERS. In light of the consistently revealed sensitivity that Stage 2 respondents had to purchase cost of IVERS, suppliers of IVERS would be most successful if they could bundle the IVERS into the price of other purchases -- most likely that of an automobile. Additionally, as most respondents with and without IVERS use experience revealed little sensitivity to usage fees, there would seem to be a market opportunity for suppliers of IVERS to derive economic return from charging for use. Indeed, this result is analogous to the demand for basic telephone service that is based on an option value (so-called "lifeline" service) rather than on use.

This marketability analysis produced a number of conclusions that can guide efforts to commercialize IVERS, but it does not address a number of institutional issues that will also affect efforts to bring these systems to market. Chapter 5 focuses on a number of these key institutional issues.

CHAPTER 5. INSTITUTIONAL ISSUES ANALYSIS

5.1 INTRODUCTION

The purpose of the Institutional Issues Analysis was to evaluate institutional issues associated with deployment of in-vehicle emergency response systems that follow the PuSHMe model, that is, that incorporate GPS technology and rely on private service centers as a first point of contact. In addition, we looked at institutional issues that impacted the PuSHMe operational test itself.

5.2 SCOPE

Because the partners performed a separate analysis covering many institutional issues, we designed our efforts to complement and expand on, rather than be duplicative of, their efforts. The partners' institutional analysis resulted in a technical memorandum *(Institutional Issues Report,* David Evans & Associates, August 1996) that explored: (1) the roles and protocols of typical Public Service Answering Points (PSAPs); (2) current existing services; and (3) legal issues for public and private service centers. In addition, the DEA memorandum presents (1) the results of a partner-led focus group on mayday technologies and services; (2) lessons from the PuSHMe Simulated Service Delivery Tests; and (3) recommendations for the implementation of a private mayday Customer Support Center (CSC).

The evaluation team's institutional analysis focused on: (1) issues related to public/private interaction, particularly those associated with the use of a privately operated CSC, and (2) institutional issues encountered in the operational test itself.

5.3 BACKGROUND

An initial assumption of both PuSHMe systems was the reliance on a privately run service center as the first line response to emergency calls. Much of our discussion of institutional issues focuses on the relationship between existing Public Service Answering Points (PSAPs) and these proposed private Customer Service Centers (CSCs).

Currently, most 911 (emergency) calls are answered and handled by PSAPs. PSAPs, sometimes known as E-911 centers, are publicly run services that respond to all calls within a given coverage area. Example PSAPs in the Puget Sound region serve counties such as King and Snohomish, as well as cities such as Kirkland and Issaquah.

CSCs are privately, rather than publicly, run. These private response centers provide services only to subscribers or paying customers. Example CSCs in the Puget Sound region include the American Automobile Association (AAA), various ambulance services, and home/office security system services. However, according to the DEA memorandum, no CSCs in the Puget Sound region currently operate as privately-run PSAPs that handle "mayday" situations. DEA describes the Mayday CSC concept as follows:

> "The mayday call would arrive with Global Positioning System (GPS) data that provides the exact location of the caller. The mayday operator would be located in a Customer Service Center (CSC) that maintains a database of customer information (e.g., medical information, emergency contacts, etc.). The CSC, operated as a subscription service, allows quick access to customer information in an emergency. The service provider is a Public Service Answering Point (PSAP), commonly known as a E-9 11 center, or another CSC that dispatches aid or communicates medical advice." ***(Institutional Issues Report,*** **David Evans & Associates, August** 1996, p.4)

In this concept, the CSC would screen calls from subscribers and then, where appropriate, call PSAPs (or another CSC) that would dispatch aid or communicate medical advice. This Mayday CSC concept, while not essential to the deployment of an IVERS, was a driving force in the way the PuSHMe test systems were established, and much of our analysis of public/private interaction focuses on this Mayday CSC concept.

5.4 METHODOLOGY

The evaluators' institutional issues analysis included: (1) a review of literature on institutional issues impacting ITS operational tests, as well as current issues surrounding 911 calls and cellular phones; (2) a review of DEA's institutional issues memorandum and the LUTE "Framework for Future Designs" report (see Appendix S); (3) participation in the partner's focus group for emergency response personnel, and (4) focused interviews with spokespersons from relevant emergency response agencies as well as with partners participating in the PuSHMe Operational Test. In addition, we shared a draft of this chapter with representatives from local PSAPs and incorporated their comments. Our interviews were structured with two aims: (1) to identify potential benefits and drawbacks of privately-run emergency response centers, and (2) to document institutional issues that impacted the operational test. Appendix T contains a copy of the letter to the PuSHMe partners outlining the institutional issues to be discussed in the interviews.

5.5 RESULTS

5.5.1 LITERATURE REVIEW

5.5.1.1 E-911 and Cellular Calls

The 911 emergency response system was first introduced in 1968, and is now the most effective way of finding help in an emergency (FCC, 1996). Currently, about 85% of 911 systems and PSAPs include some form of the upgraded Enhanced 911 (E-911) feature (FCC, 1996). E-911 enables 911 calls to automatically be routed to the most appropriate PSAP. These calls also include the caller's telephone number and the location of the telephone. In recent years, the number of 911 calls from cellular phones has been growing rapidly. In areas with E-911, cellular 911 calls, unlike landline 911 calls, present difficulties for PSAPs since they are not routed to the most appropriate PSAP and do not include the caller's telephone number and location.

Addressing the difficulties associated with cellular E-911 calls was an important motivation in the original conception of the PuSHMe Operational Test. PuSHMe addressed these difficulties through the use of (1) in-vehicle emergency response systems that utilize GPS technology to automatically locate vehicles and (2) private CSC operators who route the call to the appropriate PSAP. One of the major benefits of this approach is the ability to relay location information from cellular calls to the appropriate PSAPs, thereby resolving the current problems associated with cellular E-9 11 calls.

Since 1993 when the PuSHMe Operational Test was first conceived, other efforts have been initiated to address these issues in a different way. Administrators of PSAPs, the cellular industry, the FCC, and others have been working toward introducing new cellular 911 call requirements to better enable PSAPs to handle cellular 911 calls. On June 12, 1996,the FCC adopted a Report and Order to ensure cellular compatibility with the Enhanced 9 11 Emergency Calling System (FCC, 1996). (For the precise wording of these requirements, see Appendix U.)

One of the FCC's objectives in adopting this Report and Order was to ensure that ongoing processes are in place that will make technological advances available to 911 service providers by giving PSAP administrators (E-911 centers) the means to acquire and utilize new technologies. Towards this end, they adopted several requirements made applicable to all cellular licensees, broad band PCS licensees, and certain Specialized Mobile Radio (SMR) licensees (referred to as "Covered Carriers"). The requirements mandate that within 12 months carriers must have initiated the actions necessary to enable them to relay a caller's Automatic Number Identification (ANI) and the location of the base station or cell site receiving a 911 call to the designated PSAP. They also mandate that within 18 months after the effective date of the rules adopted, carriers must have completed these actions. These capabilities will allow the PSAP attendant to call back if the 9 11 call is disconnected.

The requirements also mandate that within five years after the effective date of the rules adopted, carriers are required to achieve the capability to identify the latitude and longitude of a mobile unit making a 911 call within a radius of no more than 125 meters in 67 percent of all cases. However, these requirements apply only if (1) a carrier receives a request for such E-9 11 services from the administrator of a PSAP that is capable of receiving and utilizing the data elements associated with the services; and (2) a mechanism for the recovery of costs relating to the provision of such services is in place.

The FCC is also seeking comments on several other related issues: (1) whether covered carriers should provide PSAPs information that locates a wireless 9 11 caller within a radius of 40 feet 90 percent of the time, using longitude, latitude, and altitude data; (2) whether wireless service providers should be required to supply location information to the PSAP regarding a 911 caller within a certain number of seconds after the 911 call is made; (3) whether wireless service providers should be required to update this location information throughout the duration of the call; and (4) what steps could be taken to enable 911 calls to be completed or serviced by mobile radio systems regardless of the availability (in the geographic area in which a mobile user seeks to place a 911 call) of the system or technology utilized by the user's wireless service.

Although requirements were set on the accuracy of the location information provided to PSAPs, no requirements were set on the location information ***technology*** that should be used. There are several different methods, in addition to the use of GPS, to determine the location of cellular callers. A number of these methods are discussed briefly in Section 1.6.

Given these changes, PSAPs should in theory be able to receive cellular 911 caller location information and telephone number within the next several years. If, in reality, PSAPs do begin receiving this information directly from cellular 911 callers, then one of the key benefits of a private Mayday CSC as conceived in PuSHMe will no longer be as critical.

5.5.1.2 Other Benefits and Issues

Even if problems associated with cellular E-911 calls disappeared, there are still numerous other potential benefits to be derived from the deployment of CSC-based IVERS such as PuSHMe. Potential benefits to the customer include: (1) the ability to signal a particular type of alert (e.g., collision vs. hijack) with the push of a button, (2) the possibility of automated alert on impact, (3) the reassurance that their chosen mayday service provider will know their exact location, (4) the customized response that a private service provider can offer, and (5) the convenience of bundled related services such as towing and repair.

In addition, private Mayday CSCs can offer potential benefits to the PSAP community. These include: (1) potentially better incident location data and vehicle tracking, (2) better personal medical histories and other personal information on those requiring assistance, (3) a screening of calls to assure they require dispatch services before being passed on to the appropriate PSAP, and (4) additional support during wide-scale emergencies (e.g., earthquake, fire, etc.).

However, there are potential drawbacks of privately-run Mayday CSCs that must be considered. These include: (1) possible additional time to handle emergency calls that are routed through CSCs to PSAPs and (2) procedural, jurisdictional, and "cultural" difference between CSCs and PSAPs. For example, PSAPs are used to direct questioning of their callers and are trained to make judgments based on this interaction. How would this fit in with the notion of a private CSC operator as intermediary?

Finally, there are social and political issues. For example, what is the appropriate level of commitment of an E-9 11 center, funded by local taxpayer dollars, in support of a private CSC, funded on a commercial basis by customers who can afford a higher grade of service than that provided by the PSAP? Will it be necessary to keep these two service centers distinct or can an appropriate working relationship be established that fairly serves all citizens?

It was potential benefits, issues, and question like these that guided our investigation of institutional issues associated with the deployment of IVERS such as PuSHMe.

5.5.2 FOCUS GROUPS

Four focus groups were scheduled by the partners and where possible the evaluation team used these as an opportunity to further explore institutional issues. The first was held on March 1996 at the University of Washington. The meeting was conducted by LUTE as part of their efforts to develop operational protocols for the CSC operators, as well as to obtain the perspective of E-91 1 professionals on institutional issues arising in the operation of a mayday system (See Appendix S). This meeting illuminated the importance of developing and maintaining a healthy relationship between public and private agencies.

At this meeting, the PSAP operators in attendance brought up a number of issues that needed to be addressed before deploying IVERS that rely on private CSCs. These included: (1) their need to hear the voice of the caller to assess each call, (2) their desire to maintain authority on any decision to dispatch public services such as police or fire personnel, (3) the need for fines on false alarms as with home security systems, (4) legal issues related to possible use of conversations or one-way listening in (i.e. the "hijack" button) as evidence in court, and (5) a general questioning of the need for private response centers at all.

A second meeting was scheduled by DEA for November but canceled on short notice due to inclement weather. Unfortunately about ten people, mostly PSAP and state patrol representatives, could not be informed and showed up anyway. This probably contributed to the fact that a substitute third meeting, held January 28th, 1997 at WSDOT's TSMC, attracted only one person who wasn't affiliated with a PuSHMe partner. Nevertheless, the January meeting was useful because Motorola brought its new partner, Bartizan Communications Inc., who put a somewhat different spin on the CSC/PSAP issue.

Bartizan discussed its plan to establish emergency service centers based on Motorola software and hardware under the name Rural/Metro Protection Services. Unlike the temporary PuSHMe CRC, Rural/Metro is a fully professional, commercial operation with backup redundancy and operators who are all licensed emergency medical technicians. Roadside services include emergency, personal security (silent alarm), directions, and polling (e.g. where's my car).

Even more significant for this discussion, part of Bartizan's roll out strategy is to convince cities and counties to outsource their emergency service operations. In this case, Rural/Metro would become the PSAP or fire department, not work with them. This would be very different than the PuSHMe CRC scenario where the private response center acts as a buffer (with both positive and negative impacts) for the public PSAP.

The January meeting also emphasized that important lessons had been learned by the partners during the course of the PuSHMe Operational Test. In the beginning, partners saw themselves as doing nearly everything that needed to be done to establish in-vehicle emergency services. By the end of the test, partners had found niches in the overall operation that they wanted to focus on. For example, both Motorola and XYPOINT decided during the test that they didn't want to operate CSCs.

Finally, the January meeting brought out one additional way that the PuSHMe operational test had been beneficial to commercial partners. Bartizan pointed out that the existence of PuSHMe was useful in their efforts to attract venture capital. Considering ITS deployment reliance on commercial products, this is an important and often overlooked benefit of operational tests.

An additional focus meeting was held in March 1997 to discuss issues related to the working relationship between private CRCs and public PSAPs. The meeting focused on possible call transfer protocols that would guide the situations which would result in a pass off from private CRC operator to public 911 operator. This emphasis on the institutional interface between public and private partners is essential if IVERS are to be successfully deployed under the PuSHMe scenario.

5.53 PARTNER INTERVIEWS

ln addition to the focus groups, interviews with the partners were used to further explore institutional issues encountered in the operational test as well as those potentially affecting future deployment. Following are highlights of these discussions.

5.5.3.1 Institutional Issues Encountered in the Operational Test

Overall, partners were extremely positive about the conduct of PuSHMe. Project goals and most partner roles were defined well early on (the role of State Patrol being the prime exception) and coordination and management by David Evans and WSDOT was highly praised.

Partners acknowledged that there were a number of challenges to be overcome, particularly early in the project, but felt all had been handled successfully. There were some early delays over resolution of legal issues required to get under contract, but this is the norm rather than the exception for any operational test involving numerous public and private partners. There were also some problems stemming from transition of people on the private side as well as the gradual withdrawal of RSPI as a partner, but once project personnel and team members were stabilized these were worked out, There was also some lack of clarity early on as to the roles of the evaluation team and the partners in the area of data gathering, but this too was worked out.

Some thought that the most regrettable operational problem was the inability to gain significant involvement from the state patrol. As an original partner, it was hoped that state patrol would provide key input and advice to the response providers, particularly in the area of institutional interaction. Unfortunately, some partners felt that the wrong person had been assigned to the project by state patrol. Whether or not that was the case, this person saw no clear role for himself. He saw PuSHMe as a DOT project with little to offer the state patrol. He therefore decided his only role was to monitor progress.

The question of whether or not a system like PuSHMe brings any benefits to the state patrol goes beyond issues concerning the operational test. It will be discussed further below along with other institutional issues affecting future deployment of a PuSHMe type emergency system

5.5.3.2 Institutional Issues Affecting Future Deployment

The PuSHMe operational test helped partners better understand the role of a private response center (PRC) in the deployment of an in-vehicle emergency response system (IVERS). Private partners felt that a PRC would be a necessary component of any early deployment scenario and pointed to existing PRCs such as those serving the Ford Lincoln RESCU system. Public partners were only slightly less certain, seeing the PRC as a viable and likely scenario, though ultimately perhaps not the only one.

Most partners see additional in-vehicle services that go beyond emergency response as being key to the relationship between PRCs and public PSAPs. They foresee the successful integration of PRCs and PSAPs as being based on the PRC handing emergency calls directly over to the PSAP while PRCs provide direct service for lower priority calls such as motorist assistance and directions. This filtering out of non-emergency calls is seen as one of the primary benefits of a PRC and the working out of protocols and data exchange for the hand off is seen as a key next activity. Many partners see this activity as parallel to the home alarm situation.

From a purely emergency response perspective, State Patrol representatives saw a PuSHMe type system to present only marginal, if any, benefits beyond that provided by cellular phones and 911 centers. They reported that with the widespread use of cellular telephones, State Patrol receives 10 to 20 calls for every breakdown on the freeway. They find that people generally know where they are and, as one person put it, “we haven’t gone searching for anybody yet.”

Patrol representatives saw the PuSHMe project as being driven by commercial concerns--something to offer people with money. They felt it was quicker to call with a cellular phone and go right into the 911 center serving that area. They wondered if the PuSHMe system didn’t provide a false sense of security and questioned if this wasn’t a lot of technology being thrown at a pretty rare situation. Finally, they wondered if adding calls from PRCs to the usual volume of calls to a 911 center didn’t degrade service to the public at the expense of serving those who can pay for a higher grade of service.

Another issue related to cellular telephones is the impact on IVERS deployment of new requirements that cellular 911 calls include automatic location and number identification (see Section 5.5.1.1). Some partners saw this as having likely impact on deployment scenarios, depending upon how the legislation was implemented. Most believed that the additional services and more accurate location (as compared to the legislative requirements) of a PuSHMe type system would still mean a profitable commercial market for a GPS and PCS based IVERS.

Some saw the cellular 911 legislation as fitting in with what was being done on the PuSHMe project. GPS would be one way for 911 centers to get the required location information and if PRCs have this plus other useful information such as traffic conditions to pass on, it might make it easier for the public and private sectors to work together.

Despite the various institutional issues that arose over the course of the project in both the operational and deployment areas, all of the partners (except State patrol) felt that their individual goals for the PuSHMe project had been achieved and that important lessons had been learned that will carry over to the next deployment phase. In particular, private sector partners learned important lessons about the significance of the institutional interface and the need to keep a dialogue going with the public agencies.

5.6 DISCUSSION AND CONCLUSIONS

PuSHMe was an operational test of a GPS-based in-vehicle emergency response system that uses a private response center for first contact and a public 911 center for emergency service dispatch. The primary institutional issues surrounding deployment of such a system involve the public/private partnership which must be evolved to make such a system feasible. These issues extend beyond protocols and guidelines which can be developed to define agreed upon interaction between the public and private sector entities. At the highest level, this partnership hinges upon assumptions about the role and cost of government and the benefits of citizenship.

The following discussion covers four high level issues: (1) the likelihood of standalone emergency services remaining a public responsibility; (2) the justification for parallel, enhanced, higher cost, private standalone emergency response services; (3) the likelihood of non-emergency in-vehicle services (e.g. directions, personal security) being combined with emergency services under the public mandate; and (4) the demand for packaged emergency and non-emergency in-vehicle services as a parallel, higher cost private response service. After discussion of these four issues, there is discussion of issues related to the effective coordination of public and private emergency response centers.

5.6.1 STANDALONE EMERGENCY SERVICES AS A PUBLIC RESPONSIBILITY

In the United States, approximately $100 billion of public revenues was spent in 1992 on roads and related services (Morris & DeCisso, 1997, pp. 13 & 14). While there is considerable debate as to whether or not road users pay the full cost of government expenditures (ibid), recent transportation reforms have explored the development of "more flexible, market-based, performance-oriented regulatory systems, incentive-based funding programs, and privatization." (Replogle, 1997) Like health care and other governmental activities that are viewed as public goods (e.g. education, law enforcement), transportation is grappling with the complex question of what level of service should be provided equally to all citizens and which services beyond that level should be provided non-uniformly on other bases such as ability/willingness to pay or to form a car pool.

Will emergency services be seen as falling above or below the line that ,separates public responsibility from private opportunity on the spectrum of possible transportation services? Different societies at different times reach different conclusions as to what constitutes an appropriate social "safety net." As a whole, French and British citizenry today demand a higher level of equitably shared government services than do the citizens of the U.S. (as can be seen in recent elections in these countries). But while the acceptance of dispensing services based on ability and willingness to pay is generally higher in the U.S. than in Europe, there are certain services where U.S. citizens have higher expectations than their European counterparts, and these include services related to individual automobile use. This is because European distances are shorter and the relationship between citizens and public transport far more positive than in the U.S., where many citizens assume the basic right to drive anywhere, anytime, and with a high degree of comfort and safety.

Taken in isolation, emergency response to automobile accidents will likely continue to be seen as falling within the spectrum of transportation services that are public responsibility. Even mountain climbers who require emergency help in national parks Iike Mount Rainier are usually provided those services at public expense (though there is debate as to the appropriateness of this). The automobile driver is engaged in a far more usual activity than mountain climbing and his or her safety is a far more general good. In addition, a public response to freeway accidents brings associated goods to the numerous other citizens and commercial vehicles traveling on the same road.

5.6.2 STANDALONE EMERGENCY SERVICES AS A PRIVATE OPPORTUNITY

Do enhancements like accurate GPS location information and special PRC knowledge about individual drivers and their health needs produce sufficiently improved emergency services to justify a parallel, higher cost private response service.3 Again looking only at emergency response, the answer is probably no. Public entities such as E-91 1 dispatchers and state patrol units are still the primary response agents and incidents in urban areas are widely reported via cellular phone. There may be a higher demand in rural areas where cellular coverage is still spotty and a disproportionate number of fatal accidents occur. However, expanding coverage for cellular phones plus future mandated location information for cellular calls to E-91 1 centers point towards a relatively low priority for standalone private emergency response services.

Some high end market for standalone private emergency response services can likely be generated through appeals to higher levels of personal security (as is some percentage of the cellular phone market), but as a standalone service, emergency response alone is probably insufficient to justify the technology and infrastructure expense of a private service.

5.6.3 INCLUDING RELATED NON-EMERGENCY IN-VEHICLE SERVICES UNDER THE PUBLIC MANDATE

If standalone emergency response services are not enough to justify the development of a private service, then the key becomes the packaging of emergency services with related non-emergency in-vehicle services. These might include automated collision notification, personal security, availability and location of desired goods and services, directions, real-time navigation, non-emergency vehicle repair, towing dispatch, and various other performance and insurance packages that would provide drivers with one-stop full coverage for any on-road problems or concerns they might have. We need to ask the same two questions about this packaged set of in-vehicle response services that we just asked about the standalone emergency service--(1) will this set of services be seen as falling under the public mandate? and (2) do these packaged services constitute a sufficiently enhanced "product" to justify a parallel, higher cost private response service?

The current trend in government reform is to reduce the level of generalized service and increase individualized options. For example, recent legislation has liberalized the use of state tolls on Federal interstates. This raises the possibility of giving drivers a choice between more congested free roads and less congested toll roads. In such a climate of reduced general service, it seems highly unlikely that enhanced non-emergency in-vehicle services such as helping drivers when they are lost will be seen as falling within the spectrum of public responsibility.

There are some signs that this climate could change. For example, the petroleum industry argues that users of transportation facilities already pay for more than the government spends on roads and related services, "concluding that road users contribute 150% of such costs to public coffers." (Morris & DeCicco, 1997, p. 13) If this argument prevails and if the cost of delivering non-emergency services continues to go down, one can imagine a strong public call for generally available additional services (i.e. why shouldn't everyone have automatic collision notification or get directions when they're lost?).

For the immediate future, however, this is a highly unlikely scenario. E-911 centers and law enforcement agents are hard pressed to handle their current volume of work, and they are unlikely to take on extra, non-emergency needs. Packaged emergency and non-emergency in-vehicle services represent a clear commercial opportunity.

5.6.4 PACKAGING EMERGENCY AND NON-EMERGENCY IN-VEHICLE SERVICES AS A PRIVATE SERVICE

From a "big picture" perspective, this leaves only the question of whether or not packaging on-road non-emergency services with emergency services produces sufficient value to justify a parallel, higher cost private response service. The answer appears to be yes, and recent activities by Ford, Baritan, and others indicate that companies and venture capitalists are willing to bet development dollars on this answer. And why not since ultimately what is being sold is a sense of security that home alarm companies, cellular phone services, and the Automobile Association of America have successfully marketed in the past to the same or similar customers. In PuSHMe and RESCU we are probably seeing the development of the enhanced, high-tech AAA of the future, perhaps combined with a car leasing and renting company as well as cellular and paging service.

Perhaps, in fact, there is sufficient value in the non-emergency services to make them a viable product independent of emergency services, which could be handled as a direct call to the E-91 1 center. This would eliminate the need to coordinate between public and private centers. For now, however, private in-vehicle service like that represented in PuSHMe are attempting to provide both non-emergency and emergency services, with emergency calls handled first by the private sector and then passed on to the public sector.

5.6.5 COORDINATING PUBLIC AND PRIVATE SECTOR PARTNERS IN THE DELIVERY OF EMERGENCY SERVICES

Within this vision of an expanded private conglomeration of future on-road services, the coordination of public and private response entities is a thorny implementation issue. How this issue is solved may go a long way towards defining the nature of future private in-vehicle response services.

Presently, two somewhat conflicting visions are coming from the private sector. On the one hand, there is a vision that focuses on non-emergency services and downplays the coordination issue. In this vision, the PRC will simply pass through any call that fits the agreed upon defiition of an emergency, perhaps with some additional location and medical information. The E-911 center will treat these calls no differently than any other call. One might just as well simply have an emergency button that dialed directly to the E-911 center. The advantages of this vision are the simplification of the institutional implementation issues and the speed with which emergency calls reach the public dispatcher.

On the other hand there is a vision where the PRCs are close partners with the E-911 centers, providing a service to the E-911 centers by pre-screening calls and providing additional valuable information as appropriate. In this vision the E-911 centers arc willing and able to work out detailed, mutually beneficial protocols and standards for the private to public hand off in return for the pre-screening and other benefits of working together. The advantages of this vision are the sharing of valuable information and, perhaps, resources, as well as the appearance to the user of a single service provider.

There are serious obstacles to this collaborative vision. In this scenario it is vital to get everybody on a common operating process, but developing common communication standards and operational protocols is a difficult task. For example, in some areas state patrol use mile markers to designate location while county police use cross streets. If PRC information is to be useful to the public dispatchers and service providers, it must fit into their language and schema. In fact, PSAPs themselves vary greatly -- from paper and pencil operations to very sophisticated map-based computer-aided dispatch, with everything in between. This is a difficult environment in which to generate general standards.

Another obstacle is a tendency among public agencies to be reluctant to add any non-public entity into their organizational procedures since it complicates their already difficult and sensitive efforts. More specifically, public dispatchers do not want private operators to make any decisions that can impact the safety of public emergency response personnel. In the near term it seems unlikely that public emergency response agencies will willingly take on private partners. More likely, things will start with a simple hand-off and, if this is successful, grow from there.

A third possibility, mentioned earlier, is that private response groups “become” the public emergency service provider through outsourcing. In this scenario, citizens who wanted increased levels of service including non-emergency services could pay additional fees and receive additional services from the same “public” provider. This scenario was not the PuSHMe model and we did not explore it in detail.

In summary, emergency on-road services will likely continue to be provided by the public sector, while private sector services will focus on various non-emergency on-road services. The difficult institutional barriers between public and private in-vehicle service providers will likely result in a very simple initial relationship where emergency services are concerned-little more than a handshake. These barriers, however, are unlikely to deter the development of private in-vehicle response centers since there will be significant demand for the enhanced, non-emergency services that they can and will provide.

CHAPTER 6. SELECTED REFERENCES

CHAPTER 1 - INTRODUCTION

Brown, Alison, and Silva, Randy, ***TIDGET™ Mayday System for Motorists,*** NAVSYS Corporation, Colorado Springs, Co., April, 1994.

Collura, John, Fisher, Donald L., Szymkowiak, Andrea, and Kaul, King, ***An Interim Report on the Evaluation of Project Northstar,*** Departments of Civil & Environmental Engineering and Industrial Engineering & Operations Research, University of Massachusetts, November, 1994.

David Evans and Associates, ***Controlled Field Testing Report, PuSHMe Operational Test - Task 2 Technical Memorandum,*** August, 1996.

Emmett, Arielle, "911 Enhanced for Cellular," ***America's NETWORK,*** April 1,1996.

Fantau, J., "Urban Positioning with GPS : A Mobile Communications Field Measurement Application," ***GPS World,*** July, 1993.

Hepsaydir, E. and Yates, W., "Performance Analysis of Mobile Positioning Using Existing CDMA Network," IEEE, 1994.

Miller, Brian, "Cellular Revolution Threatening E-911," ***Government Technology,*** October, 1994.

Press Release, Ford Motor Company, Lincoln-Mercury Division, "Lincoln RESCU: Personal Security at the Touch of a Button,"

Press Release, Ford Motor Company, Automotive Components Division, "Lincoln RESCU Passes Field Tests,"

Sweeney, L.E. Jr., Zavoli, W.B., and Loughmiller, G., "Comparitive Performance of Various Automotive Navigation Technologies," ***IEEE-IEE Vehicle Navigation &Information Systems Conference,*** **1993.**

CHAPTER 2 - PERFORMANCE ANALYSIS

Kaplan, Elliot D., Editor, ***Understanding GPS : Principles and Applications,*** Artech House, Norwood, MA, 1996

Misra, Pratap N., ***Integrated use of GPS and GLONASS in Civil Aviation,*** MIT Lincoln Laboratory Journal, Vo1 6. No. 2, Summer/Fall 1993. MIT, Lexington, MA.

CHAPTER 4 - MARKET ANALYSIS

Gautschi, David A. and Rao, Vithala R., 'A Methodology for Specification and Aggregation in Product Concept Testing', January, 1987.

CHAPTER 5 - INSTITUTIONAL ISSUES ANALYSIS

David Evans and Associates, ***Institutional Issues, PuSHMe Operational Test - Task I Technical Memorandum,*** August, 1996.

Federal Communications Commission, ***Report and Order AND Further Notice of Proposed Rulemaking: In the Matter of Revision of the Commission's Rules to Ensure Compatability with Enhanced 911 Emergency Calling Systems,*** CC Docket No. 94-102, July 26, 1996 (Adopted June 12, 1996).

Morris, Hugh & John DeCicco, "Revisiting the Extent to which User Fees Cover and Expenditures in the United States," Paper No. 97-0341, Transportation Research Board 76th Annual Meeting, 1997.

Reploge, Michael, "Overcoming Barriers to Market-based Transportation Reform," Paper No. 97-0243, Transportation Research Board 76th Annual Meeting, 1997.

SAIC, IVHS Institutional Issues and Case Studies Field Guide, Prepared for USDOT, Volpe National Transportation Systems Center (unpublished), June, 1993.

SAIC, IVHS Institutional Issues and Case Studies Analysis and Lessons Learned, Prepared for USDOT, Volpe National Transportation Systems Center, March, 1994.

APPENDICES