

Table of Contents

Preface

Abstract 3

List of Illustrations 11

List of Tables 13

Synopsis of Chapters 15

Acknowledgements 17

Author's Declaration 19

Chapter 1: Introduction

1. Aims 23

2. Rationale 23

3. Guidelines, Principles and Heuristics 24

 3.1 Guidelines 24

 3.2 Principles 25

 3.3 The Author's Principles Set 27

 3.4 Heuristics 32

 3.5 Nielsen's Heuristics 34

4. Comparison of Guidelines, Principles and Heuristics 36

 4.1 Guidelines versus Heuristics 36

 4.2 Comparing Heuristics 37

 4.3 Comparing Principles 37

5. Usability and Usability Evaluation Methods 40

 5.1 Usability 40

 5.2 Usability Evaluation Methods 42

 5.3 Heuristic Evaluation 44

 5.4 Comparative Usability Evaluation 46

 5.4.1 Summary of Related Research 46

 5.4.2 Some Conclusions from Related Research 48

5.5 Cognitive Psychology and Usability 50

Summary of Chapter 1 51

Chapter 2: Experiment 1

1. Introduction..... 53

2. The Evaluation Materials..... 54

2.1 Nielsen's Heuristics..... 54

2.2 The Principles Set (Short Version)..... 54

3. The Software..... 55

4. Method..... 58

4.1 Design..... 58

4.2 Subjects..... 58

4.3 Procedure..... 59

4.4 Pilot Study..... 60

5. Results..... 60

5.1 Problem Counts per Subject..... 60

5.2 Severity Ratings per Subject..... 62

5.3 Problem Tokens..... 63

5.3.1 Problem Distribution..... 63

5.3.2 Problem Listings..... 67

5.4 Problem Types..... 69

5.4.1 Problem Severity and Problem Types..... 69

5.4.2 Problem Distributions by Type..... 71

5.4.3 Problem Types per Subject..... 75

6. Summary of Results..... 76

7. Discussion..... 77

7.1 Severity ratings..... 78

7.2 Problem Types..... 79

7.3 Problem Overlap..... 79

Summary of Chapter 2..... 81

Chapter 3: Experiment 2

1. Introduction..... 83

2. The Evaluation Materials..... 83

2.1 Nielsen's Heuristics..... 84

2.2 The Principles Set (Full Version)..... 85

3. The Software..... 87

4. Method..... 89

4.1 Design..... 89

4.2 Subjects..... 89

4.3 Procedure..... 90

4.4 Pilot Study..... 91

5. Results..... 91

5.1 Problem Counts per Subject..... 92

5.2 Problem Tokens..... 94

5.2.1 Problem Distribution..... 94

5.2.2 Problem Listings..... 95

5.3 Problem Types..... 96

5.3.1 Problem Distributions by Type..... 96

5.3.2 Problem Types per Subject..... 97

5.3.3 Pedagogic/Non-Pedagogic Problems per Subject..... 98

6. Summary of Results..... 99

7. Discussion..... 99

7.1 Problem Counts..... 99

7.2 Problem Overlap..... 101

7.3 Problem Extraction..... 102

7.4 Problem Prediction..... 102

Summary of Chapter 3..... 103

Chapter 4: Cumulative Problem Curves	
1. Introduction	105
2. Nielsen's 'Three to Five Evaluators' Claim	105
2.1 Description	105
2.2 Validation	108
3. Cumulative Problem Curves and Probability Theory	109
4. Experiments 1 and 2: Cumulative Curves by Problem	113
5. The Problem Reduction Process	116
5.1 A Model	116
5.2 An Illustration: Experiment 2	118
6. Experiments 1 and 2: Internal Validity	121
6.1 Inter-rater Reliability	121
6.2 Cumulative Curves by Type	123
7. External Validity: Two UEM Studies	125
7.1 Nielsen (1994b)	126
7.2 Wright & Monk (1991)	128
8. Summary of Results	131
9. Discussion	132
9.1 Model and Theory	132
9.2 Problem Reduction	133
9.3 Reliability	134
Summary of Chapter 4	134
Chapter 5: Experiment 3	
1. Introduction	137
2. Reason's S-R-K (GEMS) Model	139
3. Software and Tasks	140
3.1 The Software	140
3.2 The Tasks	140
3.2.1 Skill Tasks (Tasks 1 and 2)	141
3.2.2 Skill-Rule Task Pairs (Tasks 3 and 4, 5 and 6)	141
3.2.2.1 Tasks 3 and 4	142
3.2.2.2 Tasks 5 and 6	142
3.2.3 Knowledge Task (Task 7)	142
4. The Evaluation Materials	144
5. Method	145
5.1 Design	145
5.2 Subjects	145
5.3 Procedure	146
5.4 Pilot Study	147
6. Results	148
6.1 Experimental Measures	148
6.2 Problem Counts Within Levels	149
6.3 Problem Counts Between Levels	150
6.4 Problem Distributions	152
6.5 Cumulative Curves, Detection Rates and Hit Rates	155
6.6 Problem Listings	159
6.7 High-Frequency Problems	161
6.8 High-Severity Problems	162
6.9 Inter-rater Reliability	163
7. Summary of Results	164
8. Discussion	166
8.1 The Evaluation Materials	166
8.2 The Task Levels	168
8.3 Predicted versus Observed Problems	170
Summary of Chapter 5	171

Chapter 6: Ticket Vending Machines

1. Introduction..... 173

2. The Three Machines..... 175

2.1 The Underground Machines (FFM and MFM)..... 175

2.2 The Overground Machine (QF)..... 177

3. Method..... 178

3.1 Station Locations..... 178

3.2 Procedure..... 178

3.2.1 Initial Inspection (1990)..... 178

3.2.2 Observation Sessions (1990-91 and 1996)..... 178

3.3 Error Categories..... 179

4. Results..... 180

4.1 Overall observations..... 180

4.2 Error Observations..... 180

4.2.1 Errors per User..... 180

4.2.2 Error Distributions and Cumulative Curves..... 181

4.3 Predicted Versus Observed Errors..... 183

4.4 Experiment 3: Test Condition..... 185

5. Summary of Results..... 186

6. Discussion..... 187

6.1 Vending Machines Data..... 187

6.2 Experiment 3 Test Data..... 188

Summary of Chapter 6..... 189

Chapter 7: VP-Lab

1. Introduction..... 191

2. Versions 1 and 2 Evaluations..... 191

2.1 Version 1..... 192

2.1.1 Version 1 Evaluation Results..... 192

2.2 Version 2..... 193

2.2.1 Version 2 Evaluation Results..... 194

3. Principles Usage..... 195

4. Summary of Results..... 197

5. Discussion..... 197

Summary of Chapter 7..... 198

Chapter 8: Conclusions and Discussion	
1. Introduction	199
1.1 Principles versus Heuristics	199
1.2 Cumulative Problem Curves	200
2. Discussion	201
2.1 Principles versus Heuristics	201
2.1.1 Internal Validity	202
2.1.1.1 Problem Count	202
2.1.1.2 Severity Assessment	203
2.1.1.3 Problem Recording	204
2.1.1.4 Subject Experience Questionnaire	205
2.1.1.5 Problem Types	206
2.1.2 External Validity	208
2.1.2.1 Evaluation Materials	208
2.1.2.2 Evaluators	210
2.1.2.3 Systems and Tasks	211
2.1.2.4 Empirical Support for Evaluation Materials	213
2.2 Cumulative Problem Curves	214
2.2.1 Internal Validity	215
2.2.1.1 Hit rate, Accuracy and Redundancy	215
2.2.1.2 Hit Rate, Prediction Rate and Detection Rate	216
2.2.2 External Validity	217
2.2.2.1 Observed Problems	217
2.2.2.2 Detection Rates	218
3. Final Remarks	222
Appendixes	
Appendixes A to F: Experiments 1 to 3 Materials	225
Appendix A: Nielsen's Heuristics	227
The Author's Principles Set	
Appendix B: Short Version (Experiment 1)	229
Appendix C: Preface to Full Principles Set (Experiment 2)	233
Appendix D: Full Principles Set (Experiment 2)	235
Appendix E: Subject Experience Questionnaire (Experiment 1 version)	263
Appendix F: Experiment 3 Assumptions	265
Appendixes G to J: Experiments 1 to 3 Results	267
Full Problem Listings	
Appendix G: Experiment 1 (prioritised)	269
Appendix H: Experiment 2 (prioritised)	275
Appendix I: Experiment 3 (predicted and observed, by frequency)	279
Appendix J: Experiment 2 Subject Protocol (report form)	281
Appendixes K and L: VP-Lab	283
Appendix K: Evaluation Tasks	
Version 1	285
Version 2	286
Appendix L: Evaluation Questionnaire	287
List of Selected Acronyms and Abbreviations	289
References	291

